Uncertainty Quantification of Coupled Problems with Applications to Lithium-ion Batteries

by

Mohammad Hadigol

M.Sc., University of Tehran, 2010

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Aerospace Engineering Sciences

2016

This thesis entitled:
Uncertainty Quantification of Coupled Problems with Applications to Lithium-ion Batteries
written by Mohammad Hadigol
has been approved for the Department of Aerospace Engineering Sciences

_____

Prof. Alireza Doostan

_____

Prof. Kurt Maute

Date _____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Hadigol, Mohammad (Ph.D., Aerospace Engineering Sciences)

Uncertainty Quantification of Coupled Problems with Applications to Lithium-ion Batteries

Thesis directed by Prof. Alireza Doostan

This thesis includes three main parts that are concerned with the propagation of uncertainty across high-dimensional coupled problems with applications to Lithium-ion batteries (LIBs). In all three parts, spectral methods involving polynomial chaos expansions (PCEs) are employed to quantify the effects of propagating the input uncertainties across the system.

In the first part, a stochastic model reduction approach based on low-rank separated representations is proposed for the partitioned treatment of the uncertainty space in coupled domain problems. Sequential construction of the sub-domain solutions with respect to the stochastic dimensionality of each sub-domain enabled by the classical FETI method drastically reduces the overall computational cost and provides a well suited framework for parallel computing. Two high-dimensional stochastic problems, a 2D elliptic PDE with random diffusion coefficient and a stochastic linear elasticity problem, have been considered to study the performance and accuracy of the proposed stochastic coupling approach.

A sampling-based UQ framework to study the effects of input uncertainties on the performance of LIBs using a full-order physics-based electrochemical model is presented in the second part. Developed based on sparse PCEs, the proposed UQ technique enables one to study the effects of LIB model uncertainties on the cell performance using a fairly *small* number of battery simulations. An $LiC_6/LiCoO_2$ cell with 19 random parameters has been considered to study the performance and accuracy of the proposed UQ approach. It was found that the battery discharge rate is a key factor affecting not only the performance variability of the cell, but also the determination of most important random inputs.

The third part provides a comprehensive review of the sampling techniques for the regression-based PCEs. Traditional sampling methods such Monte Carlo, Latin hypercube, quasi-Monte Carlo, optimal design of experiments, Gaussian quadratures as well as more recent techniques such as coherence-optimal and randomized quadratures are discussed. In addition, hybrid sampling methods referred to by the *alphabetic-*

coherence-optimal techniques which are a combination of the alphabetic optimality criteria and the coherence-optimal sampling method are proposed. It was observed that the alphabetic-coherence-optimal techniques outperform other sampling methods, specially when high-order PCEs are employed and/or the oversampling ratio is low.

Dedication

To my parents and my brothers, for their endless love, support and encouragement.

And to all advocates of scientific reasoning and rational thinking.

Far from being demeaning to human spiritual values, scientific rationalism is the crowning glory of the human spirit.

RICHARD DAWKINS

Acknowledgements

While my name may be alone on the front cover of this dissertation, I am by no means its sole contributor. Rather, there are a number of people behind this piece of work who deserve to be both acknowledged and thanked here.

I would like to express my special appreciation and thanks to my advisor Prof. Doostan, you have been a tremendous mentor for me. I would like to thank you for encouraging my research and for allowing me to grow as an independent researcher. Working with you was an unforgettable experience. I would also like to thank all of the current and former members of the Uncertainty Quantification group, without whom this research would not exist. I am forever indebted to all of you.

The members of my dissertation committee, Prof. Felippa, Prof. Maute, Prof. Lee, and Prof. Jones, have generously given their time and expertise to better my work. I thank them for their contribution and their good-natured support.

Lastly, a special thanks to my family. Words cannot express how grateful I am to my mother, father, and brothers for all of the sacrifices they have made on my behalf.

Contents

Chapter

Tables

Table

Figures

Figure

Chapter 1

Introduction

As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality.

ALBERT EINSTEIN

1.1 Uncertainty Sources

Simulation-based prediction of most physical systems is subject to either lack of knowledge about the governing physical laws or incomplete/limited information about model parameters such as material properties, initial, or boundary conditions. The former is referred to as epistemic uncertainty, while the later is known as aleatory uncertainty. Epistemic uncertainty can be reduced by increasing our knowledge, e.g. performing more experimental investigations and/or developing new physical models. In contrast, aleatory uncertainty cannot be reduced as it arises naturally from observations, e.g. measurements, of the system. Additional experiments can only be used to better characterize the variability. It is known that measurements of any physical quantity can never be exact and the measured value is known within a certain range of uncertainty. The measurement uncertainty is usually divided into two components: random and systematic uncertainties [103]. Unknown and unpredictable changes in the experiment cause random uncertainties while systematic uncertainties usually come from measuring instruments. Random uncertainties may be reduced by averaging over a larger number of observations, but systematic errors are very hard to detect and most of the times non-reducible. Since the measurement errors are inevitable, a reliable and more

realistic computer simulation should be able to take into account for the uncertainties in the model input parameters. Uncertainty quantification (UQ), an emerging field in computational engineering and science, is concerned with the development of rigorous and efficient solutions to this exercise. In general, UQ provides tools for assessing the credibility of model predictions and facilitating decision making under uncertainty [146, 22, 205]. UQ is also utilized for quantitative validation of simulation models [258, 148, 155] and robust design optimization under uncertainty [19, 67, 13].

Next, a brief review of different UQ techniques available in the literature in presented.

1.2 Uncertainty Quantification Techniques

A major class of UQ approaches are probabilistic where uncertain parameters are represented by random variables/processes. Non-probabilistic approaches such as fuzzy set theory [267], interval theory [164] and evidence theory [213] are more popular in areas like cognitive sciences, robotics and economics. Non-probabilistic UQ methods are not discussed in this work.

Although it is difficult to build a proper taxonomy of the UQ techniques, in general, one may classify them into two groups: intrusive and non-intrusive methods. Reformulation of the governing equations of the mathematical models describing the physical processes is needed when one employs intrusive UQ techniques. In other words, one needs to modify the existing deterministic solvers. On the other hand, non-intrusive UQ schemes are developed to facilitate the use of existing deterministic solvers and treat them as a black box. These methods work based on sampling the output quantity of interest (QoI) using these deterministic black boxes.

The most general sampling based UQ technique is the classical Monte Carlo (MC) method. MC schemes work based on generating independent samples of the random inputs according to their prescribed probability density function (PDF) and then simulating a (deterministic) black box model to obtain the realizations of the output QoI. Then, one may compute the statistics, such as mean and variance, as well as the PDF of the output QoI using sample averages from the ensemble of realizations of the output QoI. Simplicity and straightforward application, ability to use the existing solvers as black box, and their independence of the number of random inputs are the main advantages of the MC sampling method. It is known that

the main drawbacks of MC schemes are slow convergence rate and high computational cost. Numerous sampling strategies such as Latin hypercube sampling [116], Markov chain Monte Carlo sampling [98], quasi-Monte Carlo sampling [174], and importance sampling [231, 174] have been proposed to accelerate the convergence rate of the classical MC method. Although these sampling strategies have shown some promising improvements, they suffer from a limited range of applications.

Among the probabilistic UQ techniques, stochastic spectral methods [95, 258] based on polynomial chaos expansions (PCEs) [254, 34] have received special attention due to their advantages over traditional UQ techniques such as perturbation-based [134] and MC sampling methods. In particular, under certain regularity conditions, these schemes converge faster than MC sampling methods [149] and, unlike perturbation methods, are not restricted to problems with small uncertainty levels [95]. Stochastic spectral methods are based on expanding the solution of interest in PC bases. The coefficients of these expansions are then computed, for instance, via Galerkin projection [95], referred to as stochastic Galerkin (SG), or pseudo-spectral collocation [259, 153], named stochastic collocation (SC).

The standard Stochastic Galerkin projection [95] is an intrusive approach. Since the PC expansions provided an elegant mathematical framework for the UQ purposes, researchers have also developed non-intrusive PC expansion techniques [113, 70, 59]. As the PC-based techniques are the main UQ method in this work, a detailed description of these schemes is presented in the following.

1.2.1 Polynomial chaos expansion

Polynomial chaos expansion was first introduced by Wiener in 1938 [254]. It was reintroduced to the engineering field in 1991 by Ghanem and Spanos [95] for the problems with Gaussian input uncertainties and later extended for non-Gaussian random inputs by using the orthogonal polynomial sets of the Askey scheme as a basis (generalized PC expansion) [260].

Let $(\Omega, \mathcal{T}, \mathcal{P})$ be a complete probability space, where Ω is the sample set and $\mathcal{P}$ is a probability measure on the $\sigma-$field $\mathcal{T}$. Also assume that the system input uncertainty has been discretized and approximated by random variables, such that the vector $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_d) : \Omega \to \mathbb{R}^d$, $d \in \mathbb{N}$, represents the set of independent random inputs. We also assume that the PDF of random variable ξ_k is denoted by $\rho(\xi_k)$, while

$\rho(\boldsymbol{\xi})$ represents the joint PDF of $\boldsymbol{\xi}$. The truncated PC representation of a finite variance random process $u(\boldsymbol{x}_0, t_0, \boldsymbol{\xi})$ for a fixed $\boldsymbol{x}_0$ and t_0 defined on $(\Omega, \mathcal{T}, \mathcal{P})$ is given by

$$\hat{u}(\boldsymbol{x}_0, t_{,0}\, \boldsymbol{\xi}) = \sum_{\boldsymbol{i} \in \mathscr{I}_{d,p}} \alpha_{\boldsymbol{i}}(\boldsymbol{x}_0, t_0) \psi_{\boldsymbol{i}}(\boldsymbol{\xi}), \tag{1.1}$$

where $\alpha_{\boldsymbol{i}}(\boldsymbol{x}_0, t_0)$ are the deterministic coefficients and $\psi_{\boldsymbol{i}}(\boldsymbol{\xi})$ are the multivariate PC basis functions. The basis functions $\psi_{\boldsymbol{i}}(\boldsymbol{\xi})$ in (1.1) are generated from

$$\psi_{\boldsymbol{i}}(\boldsymbol{\xi}) = \prod_{k=1}^{d} \psi_{i_k}(\xi_k), \quad \boldsymbol{i} \in \mathscr{I}_{d,p}, \tag{1.2}$$

where $\psi_{i_k}(\xi_k)$, $k = 1, \ldots, d$, are univariate polynomials of degree $i_k \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$ orthogonal with respect to $\rho(\xi_k)$, i.e.,

$$\mathbb{E}[\psi_{i_k} \psi_{j_k}] = \int \psi_{i_k}(\xi_k) \psi_{j_k}(\xi_k) \rho(\xi_k) \mathrm{d}\xi_k = \delta_{i_k j_k} \mathbb{E}[\psi_{i_k}^2], \tag{1.3}$$

see, e.g., Table 1.1. Here, $\delta_{i_k j_k}$ is the Kronecker delta and $\mathbb{E}[\cdot]$ denotes the mathematical expectation operator. The multi-index $\boldsymbol{i}$ in (1.1) is $\boldsymbol{i} = (i_1, \cdots, i_d) \in \mathscr{I}_{d,p}$ and the set of multi-indices $\mathscr{I}_{d,p}$ is defined by

$$\mathscr{I}_{d,p} = \{\boldsymbol{i} = (i_1, \cdots, i_d) \in \mathbb{N}_0^d : \|\boldsymbol{i}\|_1 \leqslant p\}, \tag{1.4}$$

where $\|\cdot\|_1$ is the l_1 norm and the size of $\mathscr{I}_{d,p}$, hence the number P of PC basis functions of total order not larger than p in dimension d, is given by

$$P = |\mathscr{I}_{d,p}| = \frac{(p+d)!}{p!d!}. \tag{1.5}$$

Table 1.1: Correspondence of Wiener-Askey polynomial chaos and probability distribution of the corresponding random variables [260].

$\rho(\xi_k)$	Polynomial type	Support
Gaussian	Hermite	(-∞,+∞)
Gamma	Laguerre	(0,+∞)
Beta	Jacobi	[a,b]
Uniform	Legendre	[a,b]

Following the orthogonality of the polynomials $\psi_{i_k}(\xi_k)$ and given that the ξ_k are independent, the PC basis functions $\psi_{\boldsymbol{i}}(\boldsymbol{\xi})$ are also orthogonal, i.e., $\mathbb{E}[\psi_{\boldsymbol{i}}\psi_{\boldsymbol{j}}] = \delta_{\boldsymbol{i},\boldsymbol{j}}\mathbb{E}[\psi_{\boldsymbol{i}}^2]$. The truncated PC expansion in (1.1) converges in the mean-square sense as $p \to \infty$ when $u(\boldsymbol{x}_0, t_0, \boldsymbol{\xi})$ has finite variance and the coefficients $\alpha_{\boldsymbol{i}}(\boldsymbol{x}_0, t_0)$ are computed from the projection equation $\alpha_{\boldsymbol{i}}(\boldsymbol{x}_0, t_0) = \mathbb{E}[u(\boldsymbol{x}_0, t_0)\psi_{\boldsymbol{i}}]/\mathbb{E}[\psi_{\boldsymbol{i}}^2]$ [260]. In (1.1), we have used the notation $\hat{u}(\boldsymbol{x}_0, t_0, \boldsymbol{\xi})$ to distinguish the truncated PC approximation of $u(\boldsymbol{x}_0, t_0, \boldsymbol{\xi})$ from its true value. For the sake of brevity, we suppress the dependence of u and α on $\boldsymbol{x}_0$ and t_0. For convenience, we also index the PC basis functions by $\{1, \ldots, P\}$ so that there is a one-to-one correspondence between $\{\psi_j(\boldsymbol{\xi})\}_{j=1}^{P}$ and $\{\psi_{\boldsymbol{i}}(\boldsymbol{\xi})\}_{\boldsymbol{i}\in\mathscr{I}_{p,d}}$.

Having defined the orthogonal PC basis functions $\psi_{\boldsymbol{i}}(\boldsymbol{\xi})$, the next task in PC-based methods is to compute the coefficients of the solution expansion $\alpha_{\boldsymbol{i}}$, either intrusively or non-intrusively, which is discussed next.

1.2.1.1 Intrusive PC-based methods

In an intrusive PCE approach, dependent variables and random inputs are replaced with their PC representations in the governing equations and the resulting residual is projected onto the stochastic PC expansion bases via a Galerkin formulation. The final system of equations to be solved in an intrusive PC expansion method is P times larger than the size of the deterministic counterpart. This approach requires a thorough modification of the existing deterministic solvers, which for complex problems might be difficult and time-consuming. Intrusive SG techniques will be discussed in more details in Section 2.3.1.

1.2.1.2 Non-intrusive PC-based methods

On the other hand, non-intrusive methods facilitate the use of existing deterministic solvers and treat them as a black box. The first task is to generate a set of N realizations of $\boldsymbol{\xi}$, denoted by $\{\boldsymbol{\Xi}^{(i)}\}_{i=1}^{N}$. Next, corresponding to these samples, N realizations of QoI, $\{u(\boldsymbol{\Xi}^{(i)})\}_{i=1}^{N}$, are computed using an available deterministic solver for the problem at hand. The last step is solving for the PC coefficients using these realizations. Several methods such as least squares regression [113], pseudo-spectral collocation [258], Monte Carlo sampling [149] and compressive sampling (CS) [59] have been employed in the literature for

this purpose. In Section 3.3, we will review the least squares regression and CS methods. Optimal sampling techniques for regression-based PCEs are also discussed in Chapter 4.

1.2.1.3 Response statistics via PC expansions

Once the PC coefficients are computed, the mean, $\mu[\cdot]$, and variance, $\mathrm{var}[\cdot]$, of $\hat{u}(\boldsymbol{\xi})$ can be directly approximated by

$$\mu[\hat{u}] = \alpha_{\mathbf{0}}, \tag{1.6}$$

and

$$\mathrm{var}[\hat{u}] = \sum_{\substack{\boldsymbol{i}\in\mathscr{I}_{d,p}\\ \boldsymbol{i}\neq\mathbf{0}}} \alpha_{\boldsymbol{i}}^2. \tag{1.7}$$

1.3 Thesis Organization and Objective

This thesis contains three main parts. The first part presented in Chapter 2 is on the propagation of uncertainty across coupled domain problems with high-dimensional random inputs. The second part presented in Chapter 3 deals with UQ of Lithium-ion batteries (LIBs) and the last part provides a comprehensive review of the sampling techniques for the regression-based PCEs presented in Chapter 4. The main UQ technique employed in all three parts is the polynomial chaos expansion approach. The objectives and organization details of this thesis are discussed in the following.

1.3.1 Part one: partitioned treatment of uncertainty in coupled domain problems

Although PC-based techniques benefit from elegant mathematical analyses, e.g., formal convergence studies, they may suffer from the so-called *curse-of-dimensionality* if executed carelessly: the computational cost may grow exponentially as a function of the number of independent random inputs [258, 149, 59], and as a function of the total number of unknowns in the system. This may be a common situation when one is dealing with UQ of engineering problems involving, for instance, coupled domains or separated scales with independent and separate sources of uncertainty. In the past few years, several alternative methods

have been proposed to introduce some form of *sparsity* in an effort to counter the curse-of-dimensionality [243, 23, 177, 57, 86, 93, 59, 1, 60]. The ideas used here are based on sparse grids, element-like partitioning, ANOVA expansions, separated representations, and more general tensor decompositions. Although these techniques have been found efficient in alleviating curse-of-dimensionality, the high computing cost is still a major bottleneck in PC expansion of high-dimensional random solutions.

Of particular interest in the present study is the problem of uncertainty propagation across coupled domain problems where independent, high-dimensional random inputs are present in each sub-domain. In such cases, the solution depends on all random inputs; hence, a direct application of PC expansion may not be feasible or at least not desirable. While integration of PC expansions with standard domain decomposition (DD) techniques may partially reduce the overall computational complexity by partitioning the physical space, e.g., see [96, 49, 235], expansions (or sampling) with respect to the combined set of random inputs is still required. Instead, we propose an approach that additionally enables a *partitioned treatment* of the stochastic space; that is, the solution is computed through a sequence of approximations with respect to the random inputs associated with each individual sub-domain, and not the combined set of random inputs. To this end, we adopt a stochastic expansion based on the so-called *separated representations* and demonstrate how it can be obtained in conjunction with a DD approach. As one of the simplest cases of such coupling, we here consider a linear problem on two non-overlapping sub-domains with a common interface. We later provide a numerical experiment featuring three coupled sub-domains. Although more elaborate coupling techniques are possible, we restrict ourselves to a finite element tearing and interconnecting (FETI) approach [82] for the sake of simplicity.

Chapter 2 starts with literature review and introduction in Section 2.1. Section 2.2 summarizes the problem formulation, where we start with a general linear stochastic PDE as an abstract problem and present its coupled formulation via FETI. Following that, in Section 2.3, we review the application of the SG scheme to the original problem as well as its coupled formulation. In Section 2.4, we present our approach based on separated representations along with an extension of the FETI algorithm to solve the resulting coupled problems at the spatial level. Finally, two numerical examples, a 2D linear diffusion problem on an L-shaped domain with random diffusivity coefficient and a 2D linear elasticity problem describing the deflection

of a cantilever beam with random Young's modulus, are considered in Section 2.5. The accuracy and performance of the proposed coupling approach are demonstrated using these two examples.

1.3.2 Part two: uncertainty quantification of Lithium-ion batteries

High energy and power densities of LIBs alongside their superior safety features have made them the number one energy storage device for a wide range of electric devices from cell phones to hybrid-electric vehicles and aerospace applications [99, 61, 152, 72]. Since the launch of the first commercial LIB in 1991 [240], significant work has been dedicated to modeling [62, 225, 33, 236, 88, 64], design optimization [90, 100, 261], and discovering new materials [39, 182, 6] for LIBs.

Among the proposed models for simulation of LIBs, the most widely used one is Newman's model [171, 62, 63], which is based on the porous electrode and concentrated solution theories. This model involves material and electrochemical properties, such as porosity of electrodes, solid particle size, and diffusion coefficients, which are measured/estimated directly or indirectly via experimental techniques [207]. Measurements of most physical quantities are accompanied by uncertainty due to accuracy limitations or natural, cell-to-cell variabilities. For example, measurements of the solid particle size in the LIB electrodes typically result in distributed values for the particle size [207, 150] rather than a single deterministic value which is mostly used in the LIB simulations. Performing additional measurements may lead to a better characterization of the solid particle size, but cannot completely eliminate the irreducible uncertainties due to natural variations. Hence, a deterministic treatment of the solid particle size may result in predictions which do not agree with the experiments well. Consequently, LIB models which incorporate discrete or continuous particle size distributions have been developed [52, 167, 250, 234, 207]. Quantifying the impact of such uncertainties is essential for reliable model-based predictions, and is the focus of the emerging field of UQ in computational engineering and science.

We present a stochastic, physics-based model for LIBs in order to study the effects of parametric model uncertainties on the cell capacity, voltage, and concentrations. To this end, the proposed UQ approach, based on sparse polynomial chaos expansions, relies on a *small* number of battery simulations. Within this UQ framework, the identification of most important uncertainty sources is achieved by performing a

global sensitivity analysis via computing the so-called Sobol' indices. Such information aids in designing more efficient and targeted quality control procedures, which consequently may result in reducing the LIB production cost. An $LiC_6/LiCoO_2$ cell with 19 uncertain parameters discharged at 0.25C, 1C and 4C rates is considered to study the performance and accuracy of the proposed UQ approach. The results suggest that, for the considered cell, the battery discharge rate is a key factor affecting not only the performance variability of the cell, but also the determination of most important random inputs.

Following the literature review in Section 3.1, Section 3.2 presents the LIB model used in this study. In Section 3.3, we present our stochastic LIB modeling approach which is based on non-intrusive PC expansions. We then continue with reviewing a global sensitivity analysis approach via the Sobol' indices. Standard experimental techniques for measuring the model parameters and their underlying distributions are discussed in Section 3.5. Finally, an $LiC_6/LiCoO_2$ cell with 19 uncertain input parameters is considered as our numerical example in Section 3.6.

1.3.3 Part three: optimal sampling strategies for regression-based PCEs

As the non-institutive PCE techniques, and in particular, regression-based PCEs gain more and more popularity among researchers, a comprehensive review of the sampling techniques is provided in Chapter 4. Traditional sampling methods such Monte Carlo, Latin hypercube, quasi-Monte Carlo, optimal design of experiments (ODE), Gaussian quadratures as well as more recent techniques such as coherence-optimal and randomized quadratures are discussed. We also propose hybrid sampling methods referred to by the *alphabetic-coherence-optimal* techniques which are a combination of the alphabetic optimality criteria used in the context of ODE and the coherence-optimal sampling method. A comparison between the empirical performance of the selected sampling methods applied to three numerical examples, including high-order PCEs, high-dimensional problems, and low oversampling ratios, is presented to provide a road map for practitioners seeking the most suitable sampling approach for their problem at hand. We observed that the *alphabetic-coherence-optimal* techniques outperform other sampling methods, specially when high-order PC expansions are employed and/or the oversampling ratio is low.

Chapter 4 starts with the literature review and introduction presented in Section 4.1. Section 4.2

summarizes the problem formulation, where we start with a general stochastic PDE as an abstract problem. Following that, in Section 4.3, some definitions and preliminaries which are useful in describing the sampling techniques discussed in Section 4.4 are presented. Standard MC, asymptotic, coherence-optimal and random quadrature sampling techniques, as well as space-filling methods such as quasi-MC and LH sampling are reviewed in Section 4.4. Moreover, Section 4.4 provides a comprehensive review on the application of various alphabetic optimality criteria in the context of regression-based PCEs. Finally, three numerical examples, i.e., the recovery of randomly generated functions, a nonlinear stochastic Duffing oscillator and stochastic estimation of the remaining useful life (RUL) of Lithium-ion batteries (LIBs), are considered in Section 4.5 to study the empirical performance of the selected sampling methods.

Chapter 2

Partitioned Treatment of Uncertainty in Coupled Domain Problems

2.1 Introduction

As discussed earlier in Section 1.3.1, we employ low-rank separated representations, a.k.a. canonical decompositions or parallel factor analysis, to enable a partitioned treatment of the stochastic space. Separated representations were first introduced in [111] to represent a multi-way tensor by a finite sum of rank-one tensors. Consequently, it has been extensively applied to several areas including image compression and classification [214, 91], telecommunication [221, 140], chemometrics [26], data mining [138, 136, 245, 3], solution of operator equations [139, 20, 21, 4], among other applications. For a comprehensive review of separated representations and their applications we refer the interested reader to [137]. Approximation techniques based on separated representation of high-dimensional stochastic functions have also been recently proposed in [176, 58, 57, 177, 178, 131, 104, 60] and proven effective in reducing the issue of curse-of-dimensionality. We here adopt a special form of separated representations for the stochastic computation of coupled domain problems.

Let $(\Omega, \mathcal{T}, \mathcal{P})$ be a complete probability space where Ω is the sample set and $\mathcal{P}$ is a probability measure on the $\sigma-$field $\mathcal{T}$. Assume that the input uncertainty has been discretized and approximated by random variables, such that the vector $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_d) : \Omega \rightarrow \mathbb{R}^d$, $d \in \mathbb{N}$, represents the set of independent random inputs associated with a PDE defined on a domain $\mathcal{D} \subset \mathbb{R}^D$, $D \in \{1, 2, 3\}$. As displayed in Fig. 2.1, we consider the case in which $\mathcal{D}$ is composed of two non-overlapping sub-domains $\mathcal{D}_1$ and $\mathcal{D}_2$ ($\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$ and $\mathcal{D}_1 \cap \mathcal{D}_2 = \emptyset$) sharing an interface boundary $\Gamma_I = \partial\mathcal{D}_1 \cap \partial\mathcal{D}_2$. We further assume that the random inputs representing the uncertainty associated with $\mathcal{D}_1$ and $\mathcal{D}_2$ are independent and denoted by

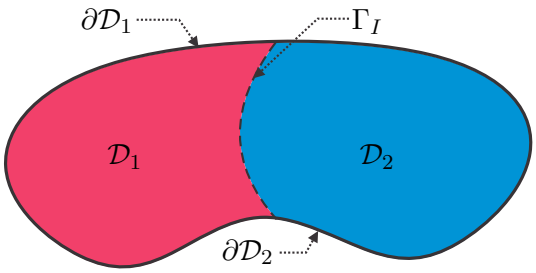


Figure 2.1: Geometry of the original problem and partitioning of $\mathcal{D}$ into two non-overlapping sub-domains $\mathcal{D}_1$ and $\mathcal{D}_2$. $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$, $\mathcal{D}_1 \cap \mathcal{D}_2 = \emptyset$, and $\partial\mathcal{D}_1 \cap \partial\mathcal{D}_2 = \Gamma_I$.

$\boldsymbol{\xi}_1 = (\xi_{1,1}, \cdots, \xi_{1,d_1}) : \Omega \to \mathbb{R}^{d_1}$, and $\boldsymbol{\xi}_2 = (\xi_{2,1}, \cdots, \xi_{2,d_2}) : \Omega \to \mathbb{R}^{d_2}$, respectively. Here, $d_1, d_2 \in \mathbb{N}$ are the sizes of $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$, respectively, and $d = d_1 + d_2$. Letting $u_1(\boldsymbol{x}, \boldsymbol{\xi}) : \bar{\mathcal{D}}_1 \times \Omega \to \mathbb{R}$ and $u_2(\boldsymbol{x}, \boldsymbol{\xi}) : \bar{\mathcal{D}}_2 \times \Omega \to \mathbb{R}$ be the sub-domain solutions, we consider the separated representation of the form

$$u_i(\boldsymbol{x}, \boldsymbol{\xi}) = \sum_{l=1}^{r} u_{0,i}^l(\boldsymbol{x}) \phi_1^l(\boldsymbol{\xi}_1) \phi_2^l(\boldsymbol{\xi}_2) + \mathcal{O}(\epsilon), \qquad i = 1, 2. \tag{2.1}$$

Here $u_{0,i}^l(\boldsymbol{x}) : \bar{\mathcal{D}}_i \to \mathbb{R}$ and $\phi_i^l(\boldsymbol{\xi}_i) : \Omega \to \mathbb{R}$, $l = 1, \ldots, r$, are, respectively, deterministic and stochastic functions – or *factors* – to be determined along with the *separation rank* r. These quantities are not fixed *a priori* and are computed through an optimization scheme such that a prescribed target accuracy ϵ is reached for a minimum r.

We note that the expansion in (2.1) is constructed by the multiplication of the unknown factors $u_{0,i}^l(\boldsymbol{x})$, $\phi_1^l(\boldsymbol{\xi}_1)$, and $\phi_2^l(\boldsymbol{\xi}_2)$; hence, finding (2.1) may be cast in the form of a non-linear problem. However, as we shall describe later, (2.1) may be computed through a sequence of alternating linear problems, where $\{u_{0,i}^l(\boldsymbol{x})\}_{l=1}^r$, $i = 1, 2$, $\{\phi_1^l(\boldsymbol{\xi}_1)\}_{l=1}^r$, or $\{\phi_2^l(\boldsymbol{\xi}_2)\}_{l=1}^r$ are solved for one at a time while others are fixed at their recent values. This alternating construction, together with the separated form of (2.1), enables computing the stochastic functions $\{\phi_1^l(\boldsymbol{\xi}_1)\}_{l=1}^r$ and $\{\phi_2^l(\boldsymbol{\xi}_2)\}_{l=1}^r$ with computational complexities that depend on d_1 and d_2 but not $d = d_1 + d_2$. This will be delineated further in Sections 2.3.2 and 2.4.4. Additionally, the representation (2.1) allows for a natural extension of DD techniques for computing $\{u_{0,i}^l(\boldsymbol{x})\}_{l=1}^r$. In the present work, we employ the standard FETI approach [82] for this purpose. Moreover, for situations where the separation rank r is small, (2.1) provides a reduced order approximation and representation of the coupled solution. We will describe the details of computing (2.1) in Section 2.4.

Among the limited earlier effort on separating random inputs for the solution of coupled problems, we particularly mention the recent work of Arnst et al. [7] that is based on the so-called *reduced chaos* expansions [229]. In particular, in each component, e.g., physics or sub-domain, the solution is expanded in a PC basis that is generated based on random inputs associated with that component. The PC coefficients are then considered to be functions of random inputs corresponding to the other component. This may be seen as a special case of (2.1). As in (2.1) we do not prescribe a stochastic basis *a priori*, our approach is different from that of [7]. The construction of (2.1) is also different from that of the reduced chaos expansions. More importantly, when the coupled solution admits a low separation rank r, the computational complexity and storage requirement of (2.1) may be significantly smaller than those of the reduced chaos expansion, especially when d is large. A brief comparison of standard PC expansions and separated representations has been presented in [60, Section 3].

Separated representations have also been utilized in the context of multi-scale problems with local uncertainties, [41]. In that work, a local-global, multi-scale iterative method has been devised in which the macro-scale global problem involves a deterministic operator with stochastic source term. The local micro-scale problems require the solution of stochastic operator equations with stochastic sources. The continuity of the solution between the global and local problems is enforced weakly via Lagrange multipliers, and separated representations have been used to solve the two problems. While for two different purposes, our coupled domain formulation overlaps with that of [41] in that we also enforce the solution continuity using Lagrange multipliers. However, several differences exist. In particular, (a) our approach is based on the formulation of a saddle-point problem as opposed to an iterative local-global method, (b) we directly approximate the solution of this saddle-point problem while in [41] a reformulation of that problem is considered, and (c) the particular choice of separated representation (2.1) and its numerical construction, both at the physical and stochastic levels, are considerably different from those adopted in [41]. An adaptation of the present work to multi-scale problems and its comparison with the approach in [41] may form an interesting future study.

2.2 Problem Formulation and Semi-discretization

2.2.1 A linear PDE with random inputs

We consider computing the solution $u(\boldsymbol{x},\boldsymbol{\xi}) : \bar{\mathcal{D}} \times \Omega \to \mathbb{R}$ satisfying the linear PDE

$$\begin{aligned} \mathcal{L}(\boldsymbol{x},\boldsymbol{\xi})(u(\boldsymbol{x},\boldsymbol{\xi})) &= f(\boldsymbol{x}) \quad \boldsymbol{x} \in \mathcal{D}, \\ \mathcal{B}(\boldsymbol{x},\boldsymbol{\xi})(u(\boldsymbol{x},\boldsymbol{\xi})) &= g(\boldsymbol{x}) \quad \boldsymbol{x} \in \partial\mathcal{D}, \end{aligned} \tag{2.2}$$

$\mathcal{P}-$ almost surely in Ω. Here, $\mathcal{L}$ is a linear differential operator and $\mathcal{B}$ is a boundary operator taking possibly various forms on different boundary segments. Additionally, as mentioned before, the uncertainty in (4.8) is assumed to have been approximated in such a way it may be represented by the random vector $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_d) : \Omega \to \mathbb{R}^d, d \in \mathbb{N}$, consisting of independent identically distributed (i.i.d.) random variables ξ_k with probability density function $\rho(\xi_k) : \mathbb{R} \to \mathbb{R}_+$. Without loss of generality, the functions $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$ in (4.8) are considered to be deterministic. We hereafter refer to (4.8) as the *original problem*.

2.2.2 Coupled formulation via domain decomposition

Given the partitioning of the physical domain $\mathcal{D}$, shown in Fig. 2.1, and consequently the random vector $\boldsymbol{\xi}$ into $\boldsymbol{\xi} = (\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$, (4.8) may be naturally cast into the non-overlapping DD formulation

$$\begin{aligned} \mathcal{L}(\boldsymbol{x},\boldsymbol{\xi}_1)(u_1(\boldsymbol{x},\boldsymbol{\xi})) &= f(\boldsymbol{x}) \quad \boldsymbol{x} \in \mathcal{D}_1, \\ \mathcal{L}(\boldsymbol{x},\boldsymbol{\xi}_2)(u_2(\boldsymbol{x},\boldsymbol{\xi})) &= f(\boldsymbol{x}) \quad \boldsymbol{x} \in \mathcal{D}_2, \\ \mathcal{B}(\boldsymbol{x},\boldsymbol{\xi}_1)(u_1(\boldsymbol{x},\boldsymbol{\xi})) &= g(\boldsymbol{x}) \quad \boldsymbol{x} \in \partial\mathcal{D}_1, \\ \mathcal{B}(\boldsymbol{x},\boldsymbol{\xi}_2)(u_2(\boldsymbol{x},\boldsymbol{\xi})) &= g(\boldsymbol{x}) \quad \boldsymbol{x} \in \partial\mathcal{D}_2, \\ u_1(\boldsymbol{x},\boldsymbol{\xi}) &= u_2(\boldsymbol{x},\boldsymbol{\xi}) \quad \boldsymbol{x} \in \Gamma_I, \end{aligned} \tag{2.3}$$

which we henceforth refer to as the *coupled problem*. Here, $u_1(\boldsymbol{x},\boldsymbol{\xi}) : \bar{\mathcal{D}}_1 \times \Omega \to \mathbb{R}$ and $u_2(\boldsymbol{x},\boldsymbol{\xi}) : \bar{\mathcal{D}}_2 \times \Omega \to \mathbb{R}$ are the solutions to (2.3). The last equation in (2.3) enforces the continuity of the solution on the interface Γ_I. Again, for the sake of simplicity, we tacitly assume that the last equation in (2.3) is sufficient to enforce that the two partial solutions $u_1(\boldsymbol{x},\boldsymbol{\xi})$ and $u_2(\boldsymbol{x},\boldsymbol{\xi})$ of (2.3) agree with the restriction of

the solution $u(\boldsymbol{x}, \boldsymbol{\xi})$ of (4.8) to the respective sub-domains. As in DD methods, our goal is then to find the sub-domain solutions $u_1(\boldsymbol{x}, \boldsymbol{\xi})$ and $u_2(\boldsymbol{x}, \boldsymbol{\xi})$ assuming independent solvers for each sub-domain.

Domain decomposition methods have been found efficient and powerful means for solving large-scale PDEs with parallel computing, [37, 191, 223, 2, 102]. In particular, non-overlapping DD methods can be classified based on the treatment of the interface solution compatibility. Primal methods, such as balancing domain decomposition (BDD) [238] and balancing domain decomposition by constraints (BDDC) [55], enforce the continuity constraint on the sub-domains interfaces via computing a unique unknown interface solution. Alternatively, in dual methods, e.g., the FETI algorithm [82], this is done by iterating on intermediary unknown variables, i.e., Lagrange multipliers. There are also hybrid methods such as dual-primal finite element tearing and interconnecting (FETI-DP) [80], that are based on the combination of these two approaches. In the present work, we use the classical FETI algorithm towards computing the coupled solution to (2.3). However, we note that our approach may also be integrated with other non-overlapping DD techniques, specially the FETI-DP.

The FETI method is based on finding the stationary or saddle point of an energy functional associated with the sub-domain solutions, which we will discuss next.

2.2.3 Semi-discrete coupled formulation

Let $\boldsymbol{u}_1(\boldsymbol{\xi}) \in \mathbb{R}^{M_1} \otimes \mathcal{W}$ and $\boldsymbol{u}_2(\boldsymbol{\xi}) \in \mathbb{R}^{M_2} \otimes \mathcal{W}$ be the vector of nodal values corresponding to the finite element (FE) discretization of u_1 and u_2 in (2.3), respectively. For simplicity, we assume that the triangulations of $\mathcal{D}_1$ and $\mathcal{D}_2$ result in matching nodes at the interface Γ_I with M_I degrees of freedom. Let also $\boldsymbol{\lambda}(\boldsymbol{\xi}) = [\lambda_1(\boldsymbol{\xi}), \cdots, \lambda_{M_I}(\boldsymbol{\xi})]^T \in \mathbb{R}^{M_I} \otimes \mathcal{W}$ be the random vector of Lagrange multipliers to enforce the solution continuity at Γ_I. Then the triple $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{\lambda})$ is the stationary point of the energy functional

$$\begin{aligned} &\pi(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{\lambda}) = \\ &\mathbb{E}\left[\frac{1}{2}\boldsymbol{u}_1^T \boldsymbol{K}_1 \boldsymbol{u}_1 - \boldsymbol{u}_1^T \boldsymbol{f}_1\right] + \mathbb{E}\left[\frac{1}{2}\boldsymbol{u}_2^T \boldsymbol{K}_2 \boldsymbol{u}_2 - \boldsymbol{u}_2^T \boldsymbol{f}_2\right] + \mathbb{E}\left[\boldsymbol{\lambda}^T \left(\boldsymbol{C}_2^T \boldsymbol{u}_2 - \boldsymbol{C}_1^T \boldsymbol{u}_1\right)\right], \end{aligned} \tag{2.4}$$

where, for $i = 1, 2$, $\boldsymbol{C}_i \in \mathbb{R}^{M_i \times M_I}$ are matrices that extract the interface nodal solutions from $\boldsymbol{u}_i$ and $\boldsymbol{K}_i(\boldsymbol{\xi}_i) : \Omega \to \mathbb{R}^{M_i \times M_i}$ and $\boldsymbol{f}_i \in \mathbb{R}^{M_i}$, are, respectively, the random stiffness matrices and deterministic

force vectors associated with the FE discretization of (2.3) on $\mathcal{D}_i$.

Assumption 1. *In the present study, we assume that $\boldsymbol{K}_1(\boldsymbol{\xi}_1)$ is symmetric positive definite, while $\boldsymbol{K}_2(\boldsymbol{\xi}_2)$ is symmetric $\mathcal{P}-$almost surely in Ω.*

Given Assumption 1, the coupled solution $\boldsymbol{u}_c = [\ \boldsymbol{u}_1^T, \boldsymbol{u}_2^T, \boldsymbol{\lambda}^T\ ]^T$ can be obtained by solving the following saddle point problem, established by the first-order optimality condition corresponding to the stationary point of π in (2.4), i.e., $\delta\pi = 0$ for arbitrary variations $\delta\boldsymbol{u}_1$, $\delta\boldsymbol{u}_2$, and $\delta\boldsymbol{\lambda}$,

$$\begin{bmatrix} \boldsymbol{K}_1(\boldsymbol{\xi}_1) & \mathbf{0} & \boldsymbol{C}_1 \\ \mathbf{0} & \boldsymbol{K}_2(\boldsymbol{\xi}_2) & -\boldsymbol{C}_2 \\ \boldsymbol{C}_1{}^T & -\boldsymbol{C}_2{}^T & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \boldsymbol{u}_1(\boldsymbol{\xi}) \\ \boldsymbol{u}_2(\boldsymbol{\xi}) \\ \boldsymbol{\lambda}(\boldsymbol{\xi}) \end{Bmatrix} = \begin{Bmatrix} \boldsymbol{f}_1 \\ \boldsymbol{f}_2 \\ \mathbf{0} \end{Bmatrix}, \quad \text{or} \quad \boldsymbol{K}_c(\boldsymbol{\xi})\boldsymbol{u}_c(\boldsymbol{\xi}) = \boldsymbol{f}_c. \tag{2.5}$$

The deterministic version of (2.5), i.e., corresponding to fixed realizations of $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$, may be solved efficiently using the FETI method, described in Section 2.4.2. This is done by eliminating $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ from (2.5) and then solving the resulting system of equations for $\boldsymbol{\lambda}$ using a projected conjugate gradient solver [82].

We next present a stochastic approximation of (4.8), when discretized by FE method in space, and (2.5) using the standard PC expansions. Following that, in Section 2.4, we introduce the proposed stochastic discretization which allows us to solve (2.5) iteratively through approximations in terms of $\boldsymbol{\xi}_1$ or $\boldsymbol{\xi}_2$.

2.3 Stochastic Discretization via Polynomial Chaos

We define the finite dimensional space

$$\mathcal{W}_p = \text{span}\{\psi_{\boldsymbol{i}}(\boldsymbol{\xi}) : \boldsymbol{i} \in \mathscr{I}_{d,p}\} \subset \mathcal{W},$$

within which we seek an approximation to a finite-variance, stochastic function $u(\boldsymbol{\xi}) \in \mathcal{W}$ by

$$u_p(\boldsymbol{\xi}) = \sum_{\boldsymbol{i} \in \mathscr{I}_{d,p}} u_{\boldsymbol{i}} \psi_{\boldsymbol{i}}(\boldsymbol{\xi}). \tag{2.6}$$

As proposed in [95], the PC coefficients $u_{\boldsymbol{i}}$ may be computed via Galerkin projection that we briefly describe next.

2.3.1 Galerkin projection

Let $\boldsymbol{u}_p(\boldsymbol{\xi}) = \sum_{\boldsymbol{i}\in\mathscr{I}_{d,p}} \boldsymbol{u_i}\psi_{\boldsymbol{i}}(\boldsymbol{\xi})$ be the PC approximation of the FE solution $\boldsymbol{u}(\boldsymbol{\xi})$ to (4.8). The coefficients $\boldsymbol{u}_i$ can be computed from the Galerkin projection

$$\mathbb{E}[\boldsymbol{v}^T(\boldsymbol{K}\boldsymbol{u}_p - \boldsymbol{f})] = 0, \qquad \forall \boldsymbol{v} \in \mathbb{R}^M \otimes \mathcal{W}_p,$$

which can be equivalently written in the form of the coupled system of equations

$$\sum_{\boldsymbol{i}\in\mathscr{I}_{d,p}} \mathbb{E}[\psi_{\boldsymbol{i}}\psi_{\boldsymbol{j}}\boldsymbol{K}]\boldsymbol{u_i} = \mathbb{E}[\psi_{\boldsymbol{j}}\boldsymbol{f}], \qquad \forall \boldsymbol{j} \in \mathscr{I}_{d,p}. \tag{2.7}$$

Following the same procedure, the Galerkin solution of the coupled problem (2.5) is computed from

$$\sum_{\boldsymbol{i}\in\mathscr{I}_{d,p}} \mathbb{E}[\psi_{\boldsymbol{i}}\psi_{\boldsymbol{j}}\boldsymbol{K}_c]\boldsymbol{u}_{c,\boldsymbol{i}} = \mathbb{E}[\psi_{\boldsymbol{j}}\boldsymbol{f}_c], \qquad \forall \boldsymbol{j} \in \mathscr{I}_{d,p}, \tag{2.8}$$

where $\boldsymbol{K}_c$ and $\boldsymbol{f}_c$ are given in (2.5), and $\boldsymbol{u}_{c,\boldsymbol{i}} = \left[\begin{array}{c} \boldsymbol{u}_{1,\boldsymbol{i}}^T, \boldsymbol{u}_{2,\boldsymbol{i}}^T, \boldsymbol{\lambda}_{\boldsymbol{i}}^T \end{array} \right]^T$. Here, $\boldsymbol{u}_{1,\boldsymbol{i}}, \boldsymbol{u}_{2,\boldsymbol{i}}$, and $\boldsymbol{\lambda_i}$ are the PC coefficients of the sub-domain solutions $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$, and the Lagrange multipliers $\boldsymbol{\lambda}$ in (2.5), respectively.

2.3.2 Need for alternative stochastic discretizations

We note that the size of the linear system of equations (2.7) and (2.8) depends on P which is related to $d = d_1 + d_2$ via (1.5). In particular, the linear system (2.7), corresponding to the PC approximation of the original problem (4.8) discretized using standard FE method with M degrees of freedom, is of size $MP \times MP$. On the other hand, assuming that $\boldsymbol{\lambda_i}$ are given, $\boldsymbol{u}_{1,\boldsymbol{i}}$ and $\boldsymbol{u}_{2,\boldsymbol{i}}$ in the coupled formulation (2.8) may be computed from linear systems of equations of sizes $M_1P \times M_1P$ and $M_2P \times M_2P$, respectively. While the latter formulation provides a reduction with respect to the size $M \approx M_1 + M_2$ of the spatial discretization, it may be computationally prohibitive when d, hence P, is large.

To alleviate this difficulty, we seek an alternative stochastic discretization that allows us to approximate the solution of the coupled problem (2.5) by solving problems in dimensions d_1 and d_2, instead of $d = d_1 + d_2$, thereby partitioning the stochastic space in addition to the physical space. To this end, we propose the approximation of the solution to (2.5) via separated representations that we describe next.

2.4 Separated Representations

Given a target accuracy ϵ, as an extension of (2.1), we consider the following approximation of the solution to (2.5),

$$\left\{ \begin{array}{c} \boldsymbol{u}_1(\boldsymbol{\xi}) \\ \boldsymbol{u}_2(\boldsymbol{\xi}) \\ \boldsymbol{\lambda}(\boldsymbol{\xi}) \end{array} \right\} = \sum_{l=1}^{r} \left\{ \begin{array}{c} \boldsymbol{u}_{0,1}^l \\ \boldsymbol{u}_{0,2}^l \\ \boldsymbol{\lambda}_0^l \end{array} \right\} \phi_1^l(\boldsymbol{\xi}_1)\phi_2^l(\boldsymbol{\xi}_2) + \mathcal{O}(\epsilon), \tag{2.9}$$

which we refer to as the separated representation with *separation rank* r. Observe that the set of all rank r approximations is not a linear subspace [104]. The deterministic vectors $\boldsymbol{u}_{0,1}^l \in \mathbb{R}^{M_1}$, $\boldsymbol{u}_{0,2}^l \in \mathbb{R}^{M_2}$, and $\boldsymbol{\lambda}_0^l \in \mathbb{R}^{M_I}$, the stochastic functions $\phi_1^l(\boldsymbol{\xi}_1) : \Omega \to \mathbb{R}$ and $\phi_2^l(\boldsymbol{\xi}_2) : \Omega \to \mathbb{R}$, as well as the separation rank r are not fixed *a priori* and are sought for to achieve the accuracy ϵ.

Due to the separated construction of the representation (2.9) with respect to variables along spatial as well as the stochastic directions $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$, the unknowns in the RHS of (2.9) may be computed from a sequence of linear approximations arising from an *alternating direction* optimization of a suitable cost function. One instance of such a multi-linear approach is described in Section 2.4.1.

The separation rank r plays an important role in the construction of separated representations. In particular, smaller accuracies ϵ entail larger r values. This, in turn, results in an increase in the computational complexity of the expansion, as we shall see in Section 2.4.1. However, when the solution to (2.5) admits a low separation rank r, (2.9) is essentially a reduced representation – or compression – of the solution. Therefore, we ideally desire to keep r as small as possible.

We note that the separated representation of the form (2.9) may not be unique [20]. However, we seek one such representation that is within an accuracy ϵ of the solution to (2.5) which will be discussed in the sequel.

Remark 1. *A more general separated approximation to the solution of (2.5) may be considered as*

$$\left\{ \begin{array}{c} \boldsymbol{u}_1(\boldsymbol{\xi}) \\ \boldsymbol{u}_2(\boldsymbol{\xi}) \\ \boldsymbol{\lambda}(\boldsymbol{\xi}) \end{array} \right\} = \sum_{l=1}^{r} \left\{ \begin{array}{c} \boldsymbol{u}_{0,1}^l \\ \boldsymbol{u}_{0,2}^l \\ \boldsymbol{\lambda}_0^l \end{array} \right\} \prod_{i=1}^{d_1} \phi^l{}_{1,i}(\xi_{1,i}) \prod_{j=1}^{d_2} \phi^l{}_{2,j}(\xi_{2,j}) + \mathcal{O}(\epsilon). \tag{2.10}$$

where ${\phi^l}_{1,i}$ and ${\phi^l}_{2,j}$ are univariate functions in $\xi_{1,i}$ and $\xi_{2,j}$, respectively. Using (2.10), one may approximate $\boldsymbol{u}_1(\boldsymbol{\xi})$, $\boldsymbol{u}_2(\boldsymbol{\xi})$, and $\boldsymbol{\lambda}(\boldsymbol{\xi})$ via a series of one-dimensional approximations with as fast as a linear increase of the computation cost with respect to d. See [20, 57] for further discussions on such growth in the case of single domain problems. This, however, is achieved by a potential increase in the separation rank r as compared to the representation (2.9). We leave further comparison of these different separated representations to a later study.

2.4.1 Alternating Rayleigh-Ritz (ARR) algorithm

To derive the separated representation (2.9), we use a Rayleigh-Ritz-type approach: we find the unknowns in the RHS of (2.9) such that they correspond to a stationary point of the energy functional π in (2.4). To do so, we adopt an alternating direction optimization approach, hence the name Alternating Rayleigh-Ritz (ARR), in order to compute (2.9) using a sequence of linear, instead of nonlinear, problems of smaller size. Notice that the stationary point of π is a saddle point, i.e, it is a minimum of π with respect to $u_1(\boldsymbol{\xi})$ and $u_2(\boldsymbol{\xi})$ and simultaneously a maximum of π with respect to $\boldsymbol{\lambda}(\boldsymbol{\xi})$.

The ARR approach consists of a sequence of stationary point iterations that are meant to converge to the saddle point of π as demonstrated in Proposition 1. Given a value for r and approximations (guesses) $\{\phi_1^l\}_{l=1}^r$ and $\{\phi_2^l\}_{l=1}^r$, the ARR iteration starts by freezing the variables along $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ directions, i.e., $\{\phi_1^l\}_{l=1}^r$ and $\{\phi_2^l\}_{l=1}^r$, and solving for the deterministic vectors $\{\boldsymbol{u}_{0,1}^l\}_{l=1}^r$, $\{\boldsymbol{u}_{0,2}^l\}_{l=1}^r$, and $\{\boldsymbol{\lambda}_0^l\}_{l=1}^r$ as the saddle point of π. We call this step a *deterministic update*. We then alternate to stochastic directions $\boldsymbol{\xi}_1$ and update $\{\phi_1^l\}_{l=1}^r$ by minimizing π, while $\{\boldsymbol{u}_{0,1}^l\}_{l=1}^r$, $\{\boldsymbol{u}_{0,2}^l\}_{l=1}^r$ and $\{\phi_2^l\}_{l=1}^r$ are fixed at their latest values. We dub this step a *stochastic update*. Similarly, we alternate to update $\{\phi_2^l\}_{l=1}^r$ by minimizing π, while the rest of the variables are fixed at their current values. This completes one full sweep of the ARR algorithm. As we initialized $\{\phi_1^l\}_{l=1}^r$ and $\{\phi_2^l\}_{l=1}^r$ arbitrarily, we repeat the ARR sweeps until the value of π does not change *much*. We then increase the separation rank r and continue the ARR sweeps until π does not change beyond some tolerance.

As we ideally want the separation rank r to be as small as possible, we may start the ARR updates

with a small r, e.g., $r = 1$ and increase r until, for instance, the maximum norm of sub-domain residuals

$$\epsilon_{res}^r = \max\left(\frac{\|\boldsymbol{f}_1 - \boldsymbol{K}_1\boldsymbol{u}_1^r\|_{L_2(\mathbb{R}^{M_1})\otimes\mathcal{W}}}{\|\boldsymbol{f}_1\|_{L_2(\mathbb{R}^{M_1})\otimes\mathcal{W}}}, \frac{\|\boldsymbol{f}_2 - \boldsymbol{K}_2\boldsymbol{u}_2^r\|_{L_2(\mathbb{R}^{M_2})\otimes\mathcal{W}}}{\|\boldsymbol{f}_2\|_{L_2(\mathbb{R}^{M_2})\otimes\mathcal{W}}}\right) \tag{2.11}$$

is below a prescribed accuracy ϵ. Here, for $i = 1, 2$, $\boldsymbol{u}_i^r = \sum_{l=1}^r \boldsymbol{u}_{0,i}^l \phi_1^l \phi_2^l$, and $\|\cdot\|_{L_2(\mathbb{R}^M)\otimes\mathcal{W}}$ is defined by $\|\boldsymbol{u}\|^2_{L_2(\mathbb{R}^M)\otimes\mathcal{W}} = \mathbb{E}[\boldsymbol{u}^T\boldsymbol{u}]$. We note that higher accuracies may be achieved by requiring smaller ϵ_{res}^r in (2.11), which in turn may necessitate higher separation ranks r.

In the following, we describe the ARR algorithm in more details.

Deterministic updates for $\{\boldsymbol{u}_{0,1}^l\}_{l=1}^r$, $\{\boldsymbol{u}_{0,2}^l\}_{l=1}^r$, and $\{\boldsymbol{\lambda}_0^l\}_{l=1}^r$. Assuming an initialization of the separation rank r, and the stochastic functions $\{\phi_1^l\}_{l=1}^r$ and $\{\phi_2^l\}_{l=1}^r$, the first step of the ARR algorithm is to update deterministic vectors $\{\boldsymbol{u}_{0,1}^l\}_{l=1}^r$, $\{\boldsymbol{u}_{0,2}^l\}_{l=1}^r$, and $\{\boldsymbol{\lambda}_0^l\}_{l=1}^r$ while freezing all other variables. Plugging (2.9) in (2.4) and enforcing the condition

$$\delta\pi = 0, \qquad \forall\, \delta\boldsymbol{u}_{0,1}^l, \delta\boldsymbol{u}_{0,2}^l, \delta\boldsymbol{\lambda}_0^l,$$

for $l = 1, \ldots, r$ and arbitrary (but consistent) variations $\delta\boldsymbol{u}_{0,1}^l$, $\delta\boldsymbol{u}_{0,2}^l$, and $\delta\boldsymbol{\lambda}_0^l$, we arrive at the saddle point problem

$$\begin{bmatrix} \hat{\boldsymbol{K}}_1 & \boldsymbol{0} & -\hat{\boldsymbol{C}}_1 \\ \boldsymbol{0} & \hat{\boldsymbol{K}}_2 & \hat{\boldsymbol{C}}_2 \\ -\hat{\boldsymbol{C}}_1^T & \hat{\boldsymbol{C}}_2^T & \boldsymbol{0} \end{bmatrix} \begin{Bmatrix} \hat{\boldsymbol{u}}_{0,1} \\ \hat{\boldsymbol{u}}_{0,2} \\ \hat{\boldsymbol{\lambda}}_0 \end{Bmatrix} = \begin{Bmatrix} \hat{\boldsymbol{f}}_1 \\ \hat{\boldsymbol{f}}_2 \\ \boldsymbol{0} \end{Bmatrix}. \tag{2.12}$$

Each entry of the linear system (2.12) has a block structure given by

$$\begin{aligned} \hat{\boldsymbol{K}}_i(l, l') &= \mathbb{E}\left[\phi_1^l \phi_2^l \boldsymbol{K}_i \phi_1^{l'} \phi_2^{l'}\right], \\ \hat{\boldsymbol{C}}_i(l, l') &= \mathbb{E}\left[\phi_1^l \phi_2^l \phi_1^{l'} \phi_2^{l'}\right] \boldsymbol{C}_i, \\ \hat{\boldsymbol{u}}_{0,i}(l) &= \boldsymbol{u}_{0,i}^l, \\ \hat{\boldsymbol{\lambda}}_0(l) &= \boldsymbol{\lambda}_0^l, \\ \hat{\boldsymbol{f}}_i(l) &= \mathbb{E}\left[\phi_1^l \phi_2^l\right] \boldsymbol{f}_i, \end{aligned} \tag{2.13}$$

for $l, l' = 1, \ldots, r$ and $i = 1, 2$. A detailed derivation of (2.12) and (2.13) can be found in A.1.

It should be noted that the form of (2.12) is similar to that of the saddle point system associated with the deterministic version of the coupled problem (2.5), except that each entry in (2.12) has a block structure, and for $r = 1$ the form actually coincides with the deterministic version. Such similarity, therefore, suggests an extension of the original FETI method [82] for a parallel and partitioned solution of (2.12). This will be discussed in more details in Section 2.4.2.

Stochastic updates for $\{\phi_1^l\}_{l=1}^r$. We freeze $\{\boldsymbol{u}_{0,1}^l\}_{l=1}^r$, $\{\boldsymbol{u}_{0,2}^l\}_{l=1}^r$, and $\{\boldsymbol{\lambda}_0^l\}_{l=1}^r$ to their updated values from (2.12). We also fix the variables $\{\phi_2^l\}_{l=1}^r$ and solve for $\{\phi_1^l\}_{l=1}^r$ by minimizing π, that is, for $l = 1, \ldots, r$, we require

$$\delta\pi = 0, \qquad \forall\ \delta\phi_1^l(\boldsymbol{\xi}_1). \tag{2.14}$$

This leads to the linear system

$$\boldsymbol{A}_1(\boldsymbol{\xi}_1)\boldsymbol{\phi}_1(\boldsymbol{\xi}_1) = \boldsymbol{b}_1, \tag{2.15}$$

where $\boldsymbol{A}_1 \in \mathbb{R}^{r\times r}$, $\boldsymbol{\phi}_1 \in \mathbb{R}^r$, and $\boldsymbol{b}_1 \in \mathbb{R}^r$ are given by

$$\begin{aligned}
\boldsymbol{A}_1(l,l') &= \boldsymbol{u}_{0,1}^{l^T}\boldsymbol{K}_1(\boldsymbol{\xi}_1)\boldsymbol{u}_{0,1}^{l'}\mathbb{E}_{\boldsymbol{\xi}_2}\left[\phi_2^l\phi_2^{l'}\right] + \boldsymbol{u}_{0,2}^{l^T}\mathbb{E}_{\boldsymbol{\xi}_2}\left[\boldsymbol{K}_2\phi_2^l\phi_2^{l'}\right]\boldsymbol{u}_{0,2}^{l'},\\
\boldsymbol{\phi}_1(l) &= \phi_1^l(\boldsymbol{\xi}_1),\\
\boldsymbol{b}_1(l) &= \boldsymbol{u}_{0,1}^{l^T}\boldsymbol{f}_1\mathbb{E}_{\boldsymbol{\xi}_2}\left[\phi_2^l\right] + \boldsymbol{u}_{0,2}^{l^T}\boldsymbol{f}_2\mathbb{E}_{\boldsymbol{\xi}_2}\left[\phi_2^l\right],
\end{aligned} \tag{2.16}$$

and $l, l' = 1, \ldots, r$. Here, $\mathbb{E}_{\boldsymbol{\xi}_2}[\cdot]$ is the mathematical expectation operator with respect to the probability density function of $\boldsymbol{\xi}_2$. A.2 provides further details on the derivation of (2.15) and (2.16).

Stochastic updates for $\{\phi_2^l\}_{l=1}^r$. Following the same procedure for $\{\phi_1^l\}_{l=1}^r$, $\{\phi_2^l\}_{l=1}^r$ are updated by solving the linear system of size r

$$\boldsymbol{A}_2(\boldsymbol{\xi}_2)\boldsymbol{\phi}_2(\boldsymbol{\xi}_2) = \boldsymbol{b}_2, \tag{2.17}$$

where, for $l, l' = 1, \ldots, r$,

$$\begin{aligned}
\boldsymbol{A}_2(l,l') &= \boldsymbol{u}_{0,1}^{l^T} \mathbb{E}_{\boldsymbol{\xi}_1}\left[\boldsymbol{K}_1 \phi_1^l \phi_1^{l'}\right] \boldsymbol{u}_{0,1}^{l'} + \boldsymbol{u}_{0,2}^{l^T} \boldsymbol{K}_2(\boldsymbol{\xi}_2) \boldsymbol{u}_{0,2}^{l'} \mathbb{E}_{\boldsymbol{\xi}_1}\left[\phi_1^l \phi_1^{l'}\right], \\
\boldsymbol{\phi}_2(l) &= \phi_2^l(\boldsymbol{\xi}_2), \\
\boldsymbol{b}_2(l) &= \boldsymbol{u}_{0,1}^{l^T} \boldsymbol{f}_1 \mathbb{E}_{\boldsymbol{\xi}_1}\left[\phi_1^l\right] + \boldsymbol{u}_{0,2}^{l^T} \boldsymbol{f}_2 \mathbb{E}_{\boldsymbol{\xi}_1}\left[\phi_1^l\right],
\end{aligned} \tag{2.18}$$

and $\mathbb{E}_{\boldsymbol{\xi}_1}[\cdot]$ is the mathematical expectation operator with respect to the probability density of $\boldsymbol{\xi}_1$.

We note that the update equations (2.15) and (2.17) only depend on $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$, respectively, and not $\boldsymbol{\xi} = (\boldsymbol{\xi}_1, \boldsymbol{\xi}_2)$, thus yielding a partitioning of the stochastic space. Approximations to the solutions of the stochastic equations (2.15) and (2.17) may be obtained using, for instance, the SG [95] or the stochastic collocation [153, 259] technique. In the present study we employ SG, described in Section 2.3, for this purpose. Specifically, for (2.15) or (2.17), we consider the PC expansions

$$\boldsymbol{\phi}_i(\boldsymbol{\xi}_i) \approx \sum_{\boldsymbol{j} \in \mathscr{I}_{d_i,p}} \boldsymbol{\phi}_{i,\boldsymbol{j}} \psi_{\boldsymbol{j}}(\boldsymbol{\xi}_i), \qquad i = 1, 2,$$

where the coefficient vectors $\{\boldsymbol{\phi}_{1,\boldsymbol{j}}\}_{\boldsymbol{j} \in \mathscr{I}_{d_1,p}}$ and $\{\boldsymbol{\phi}_{2,\boldsymbol{j}}\}_{\boldsymbol{j} \in \mathscr{I}_{d_2,p}}$ are computed via Galerkin projection. Let P_1 and P_2 be the size of $\mathscr{I}_{d_1,p}$ and $\mathscr{I}_{d_2,p}$ evaluated from (1.5) by setting $d = d_1$ and $d = d_2$, respectively. Then, the Galerkin systems to be solved for $\{\boldsymbol{\phi}_{1,\boldsymbol{j}}\}_{\boldsymbol{j} \in \mathscr{I}_{d_1,p}}$ and $\{\boldsymbol{\phi}_{2,\boldsymbol{j}}\}_{\boldsymbol{j} \in \mathscr{I}_{d_2,p}}$ are of size $rP_1 \times rP_1$ and $rP_2 \times rP_2$, respectively. Notice that, for moderate values of separation rank r, the sizes of these linear systems may be significantly smaller than those of (2.7), particularly when d_1 and d_2 are large. Such a reduction is due to the use of separated representation (2.9) along with the ARR algorithm for its construction.

We next demonstrate that the sequence of separated approximations of the form (2.9) generated by the ARR algorithm results in a non-increasing sequence of π in $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$, and a non-decreasing sequence of π in $\boldsymbol{\lambda}$. In other words, the ARR algorithm iteratively improves the approximation to the saddle point problem (2.5) unless $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, or $\boldsymbol{\lambda}$ does not change throughout the iterations. In the latter scenario, either the separation rank r needs to be increased or the separated approximation has converged to the solution of (2.5).

Proposition 1 (Iterative improvement of separated representation)**.** *Let* $(\boldsymbol{u}_1^{(k)}, \boldsymbol{u}_2^{(k)}, \boldsymbol{\lambda}^{(k)})$ *and* $(\boldsymbol{u}_1^{(k+1)}, \boldsymbol{u}_2^{(k+1)}, \boldsymbol{\lambda}^{(k+1)})$ *be the rank* r *separated representation of the solution* $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{\lambda})$ *after* k *and* $k+1$*, respectively, ARR up-*

dates (2.12), (2.15), and (2.17). Then,

$$\pi(\boldsymbol{u}_1^{(k+1)}, \boldsymbol{u}_2^{(k+1)}, \boldsymbol{\lambda}^{(k)}) \leq \pi(\boldsymbol{u}_1^{(k)}, \boldsymbol{u}_2^{(k)}, \boldsymbol{\lambda}^{(k)}) \leq \pi(\boldsymbol{u}_1^{(k)}, \boldsymbol{u}_2^{(k)}, \boldsymbol{\lambda}^{(k+1)}). \tag{2.19}$$

Proof. We first note that the updates (2.15) and (2.17) along $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ directions, respectively, lead only to minimizations of π, see A.2. Additionally, $\hat{\boldsymbol{u}}_{0,1}$ and $\hat{\boldsymbol{u}}_{0,2}$ (or equivalently, $\{\boldsymbol{u}_{0,1}^l\}_{l=1}^r$ and $\{\boldsymbol{u}_{0,2}^l\}_{l=1}^r$) in (2.12) are obtained by minimizing π, thus implying the left inequality in (2.19). Finally, the right inequality in (2.19) holds as π may increase only through the updates $\hat{\boldsymbol{\lambda}}_0$ in (2.12). □

The ARR algorithm is summarized in Algorithm 1.

Algorithm 1 Alternating Rayleigh-Ritz (ARR) Algorithm

- **Input**: target accuracy ϵ
- **Output**: separation rank r, $\{\boldsymbol{u}_{0,1}^l\}_{l=1}^r$, $\{\boldsymbol{u}_{0,2}^l\}_{l=1}^r$, $\{\boldsymbol{\lambda}_0^l\}_{l=1}^r$, $\{\phi_1^l\}_{l=1}^r$, and $\{\phi_2^l\}_{l=1}^r$
- Set $r = 1$; (randomly) initialize ϕ_1^1 and ϕ_2^1

while $\epsilon_{res}^r > \epsilon$ **do**

  while π changes more than some tolerance tol_r **do**

    - Update $\{\boldsymbol{u}_{0,1}^l\}_{l=1}^r$, $\{\boldsymbol{u}_{0,2}^l\}_{l=1}^r$, and $\{\boldsymbol{\lambda}_0^l\}_{l=1}^r$ by solving (2.12)
    - Update $\{\phi_1^l\}_{l=1}^r$ by solving (2.15)
    - Update $\{\phi_2^l\}_{l=1}^r$ by solving (2.17)

  end while

  - Compute the residual error ϵ_{res}^r
  - Set $r = r + 1$; (randomly) initialize ϕ_1^r and ϕ_2^r

end while

Remark 2. *It is worthwhile noting that the ARR algorithm presented here may be interpreted as a variation of the block Gauss-Seidel or multiplicative Schwarz [223] methods.*

To complete our discussion on the ARR algorithm we next present an extension of the classical FETI algorithm to solve the saddle point problem (2.12).

2.4.2 FETI solution of saddle point problem (2.12)

We are interested in parallel solution of (2.12) such that minimal exchange of information between the sub-domain solvers is required. This may be achieved by, for instance, an extension of the classical

FETI method for the particular case of problem (2.12). This is the approach we take in the present study. FETI, one of the widely used DD approaches, was first introduced by Farhat and Roux [82] for the parallel FE solution of second order, self-adjoint elliptic equations. It has also been successfully extended to many other problems; multiple right-hand sides [89], transient problems [79], plate bending problems [81], and non-linear problems [94]. We stress that, for the sake of simplicity, we use the original FETI [82] approach, but with some extensions one may also utilize more recent versions of FETI such as FETI-DP [80].

The FETI method is a non-overlapping DD scheme based on partitioning the computational domain into a number of sub-domains and enforcing the continuity of the solution across the sub-domain interfaces via Lagrange multipliers. In practice, after partitioning the original domain $\mathcal{D}$, there might be some sub-domains with insufficient Dirichlet boundary conditions resulting in positive semi-definite sub-domain stiffness matrices. These are also known as floating sub-domains. For a partition consisting of two sub-domains $\mathcal{D}_1$ and $\mathcal{D}_2$, we assume that $\mathcal{D}_2$ has no Dirichlet boundary condition and is a *floating* sub-domain with symmetric positive semi-definite stiffness matrix $\boldsymbol{K}_2(\boldsymbol{\xi}_2)$, while $\mathcal{D}_1$ is non-floating with symmetric positive definite stiffness matrix $\boldsymbol{K}_1(\boldsymbol{\xi}_1)$. Therefore, the second equation in (2.12), i.e.,

$$\hat{\boldsymbol{K}}_2 \hat{\boldsymbol{u}}_{0,2} = \hat{\boldsymbol{f}}_2 - \hat{\boldsymbol{C}}_2 \hat{\boldsymbol{\lambda}}_0, \tag{2.20}$$

is solvable if and only if, [82],

$$\hat{\boldsymbol{R}}_2 \left(\hat{\boldsymbol{f}}_2 - \hat{\boldsymbol{C}}_2 \hat{\boldsymbol{\lambda}}_0 \right) = \boldsymbol{0}, \tag{2.21}$$

where $\hat{\boldsymbol{R}}_2$ spans the null space of $\hat{\boldsymbol{K}}_2$, i.e., $\mathrm{range}(\hat{\boldsymbol{R}}_2) = \ker(\hat{\boldsymbol{K}}_2)$. In Section 2.4.2.1 we will discuss how $\hat{\boldsymbol{R}}_2$ is set in our formulation. The solution of (2.12) can then be written as

$$\begin{aligned}
&\left(\hat{\boldsymbol{C}}_1^T \hat{\boldsymbol{K}}_1^{-1} \hat{\boldsymbol{C}}_1 + \hat{\boldsymbol{C}}_2^T \hat{\boldsymbol{K}}_2^{+} \hat{\boldsymbol{C}}_2 \right) \hat{\boldsymbol{\lambda}}_0 = -\hat{\boldsymbol{C}}_1^T \hat{\boldsymbol{K}}_1^{-1} \hat{\boldsymbol{f}}_1 + \hat{\boldsymbol{C}}_2^T \left(\hat{\boldsymbol{K}}_2^{+} \hat{\boldsymbol{f}}_2 + \hat{\boldsymbol{R}}_2 \hat{\boldsymbol{\alpha}}_0 \right), \\
&\hat{\boldsymbol{u}}_{0,1} = \hat{\boldsymbol{K}}_1^{-1} \left(\hat{\boldsymbol{f}}_1 - \hat{\boldsymbol{C}}_1 \hat{\boldsymbol{\lambda}}_0 \right), \\
&\hat{\boldsymbol{u}}_{0,2} = \hat{\boldsymbol{K}}_2^{+} \left(\hat{\boldsymbol{f}}_2 - \hat{\boldsymbol{C}}_2 \hat{\boldsymbol{\lambda}}_0 \right) + \hat{\boldsymbol{R}}_2 \hat{\boldsymbol{\alpha}}_0,
\end{aligned} \tag{2.22}$$

where $\hat{\boldsymbol{K}}_1^{-1}$ and $\hat{\boldsymbol{K}}_2^{+}$ are the inverse and pseudo-inverse of $\hat{\boldsymbol{K}}_1$ and $\hat{\boldsymbol{K}}_2$, respectively. Here, $\hat{\boldsymbol{\alpha}}_0$ is an unknown vector of coefficients. As it can be seen, there are four unknowns in (2.22) while only three equations are available. The orthogonality condition in (2.21) is the fourth equation which makes (2.22)

determined. We can write the FETI *interface problem* by combining (2.21) and the first equation in (2.22) as

$$\begin{bmatrix} \hat{\boldsymbol{F}}_I & -\hat{\boldsymbol{R}}_2^I \\ -\hat{\boldsymbol{R}}_2^{I^T} & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \hat{\boldsymbol{\lambda}} \\ \hat{\boldsymbol{\alpha}}_0 \end{Bmatrix} = \begin{Bmatrix} \hat{\boldsymbol{d}} \\ -\hat{\boldsymbol{e}} \end{Bmatrix}, \tag{2.23}$$

where

$$\begin{aligned} \hat{\boldsymbol{F}}_I &= \hat{\boldsymbol{C}}_1^T \hat{\boldsymbol{K}}_1^{-1} \hat{\boldsymbol{C}}_1 + \hat{\boldsymbol{C}}_2^T \hat{\boldsymbol{K}}_2^{+} \hat{\boldsymbol{C}}_2, \\ \hat{\boldsymbol{R}}_2^I &= \hat{\boldsymbol{C}}_2^T \hat{\boldsymbol{R}}_2, \\ \hat{\boldsymbol{d}} &= \hat{\boldsymbol{C}}_2^T \hat{\boldsymbol{K}}_2^{+} \hat{\boldsymbol{f}}_2 - \hat{\boldsymbol{C}}_1^T \hat{\boldsymbol{K}}_1^{-1} \hat{\boldsymbol{f}}_1, \\ \hat{\boldsymbol{e}} &= \hat{\boldsymbol{R}}_2^T \hat{\boldsymbol{f}}_2. \end{aligned} \tag{2.24}$$

The solution to (2.12) can then be obtained by first solving the interface problem (2.23) for $\hat{\boldsymbol{\lambda}}$ and $\hat{\boldsymbol{\alpha}}_0$ and then by substituting these into the last two equations in (2.22) to compute $\hat{\boldsymbol{u}}_{0,1}$ and $\hat{\boldsymbol{u}}_{0,2}$. In order to avoid an explicit assembly of $\hat{\boldsymbol{F}}_I$ in (2.23), the preconditioned conjugate projected gradient (PCPG) algorithm of [82], which utilizes matrix-vector products, may be extended to compute $\hat{\boldsymbol{\lambda}}$ in (2.23). Algorithm 2 summarizes the steps of the PCPG solver, in which ϵ_{PCPG} is a stopping criterion, $\hat{\boldsymbol{P}} = \boldsymbol{I} - \hat{\boldsymbol{R}}_2^I (\hat{\boldsymbol{R}}_2^{I^T} \hat{\boldsymbol{R}}_2^I)^{-1} \hat{\boldsymbol{R}}_2^{I^T}$ is an orthogonal projection onto the null space of $\hat{\boldsymbol{R}}_2^I$, $\boldsymbol{I}$ is the identity matrix of size $rM_I \times rM_I$, and $\bar{\boldsymbol{F}}_I^{-1}$ is a suitable preconditioner. For a detailed description of the PCPG solver and different choices of preconditioners we refer the interested reader to [197, 133, 244, 193, 92, 38]. In the numerical result of Section 2.5.2, we have adopted the preconditioner $\bar{\boldsymbol{F}}_I^{-1}$ proposed in [133]. An *optimal* selection or design of $\bar{\boldsymbol{F}}_I^{-1}$, however, requires further study.

After computing $\hat{\boldsymbol{\lambda}}$, we obtain $\hat{\boldsymbol{\alpha}}_0$ from

$$\hat{\boldsymbol{\alpha}}_0 = (\hat{\boldsymbol{R}}_2^{I^T} \hat{\boldsymbol{R}}_2^I)^{-1} \hat{\boldsymbol{R}}_2^{I^T} (\hat{\boldsymbol{F}}_I \hat{\boldsymbol{\lambda}} - \hat{\boldsymbol{d}}).$$

Next, we propose a method to set $\hat{\boldsymbol{R}}_2$ which does not require an explicit computation of the null space of $\hat{\boldsymbol{K}}_2$.

Algorithm 2 The FETI PCPG Algorithm

• Initialize $\hat{\boldsymbol{\lambda}}_0 = \hat{\boldsymbol{R}}_2^I(\hat{\boldsymbol{R}}_2^{I^T}\hat{\boldsymbol{R}}_2^I)^{-1}\hat{\boldsymbol{R}}_2^T\hat{\boldsymbol{f}}_2$, $\boldsymbol{w}_0 = \hat{\boldsymbol{P}}\hat{\boldsymbol{d}} - \hat{\boldsymbol{P}}\hat{\boldsymbol{F}}_I\hat{\boldsymbol{\lambda}}_0$, $k = 1$
while $||\boldsymbol{w}_{k-1}||_{L_2}/||\hat{\boldsymbol{d}}||_{L_2} \geqslant \epsilon_{PCPG}$ **do**
  • $\boldsymbol{z}_{k-1} = \bar{\boldsymbol{F}}_I^{-1}(\hat{\boldsymbol{P}}^T\boldsymbol{w}_{k-1})$
  • $\boldsymbol{y}_{k-1} = \hat{\boldsymbol{P}}\boldsymbol{z}_{k-1}$
  • $s_k = \boldsymbol{y}_{k-1}^T\boldsymbol{w}_{k-1}/\boldsymbol{y}_{k-2}^T\boldsymbol{w}_{k-2}$ $\quad (s_1 = 0)$
  • $\boldsymbol{p}_k = \boldsymbol{y}_{k-1} + s_k\boldsymbol{p}_{k-1}$ $\quad (\boldsymbol{p}_1 = \boldsymbol{y}_0)$
  • $\gamma_k = \boldsymbol{y}_{k-1}^T\boldsymbol{w}_{k-1}/\boldsymbol{p}_k^T\hat{\boldsymbol{F}}_I\boldsymbol{p}_k$
  • $\hat{\boldsymbol{\lambda}}_k = \hat{\boldsymbol{\lambda}}_{k-1} + \gamma_k\boldsymbol{p}_k$
  • $\boldsymbol{w}_k = \boldsymbol{w}_{k-1} - \gamma_k\hat{\boldsymbol{P}}\hat{\boldsymbol{F}}_I\boldsymbol{p}_k$
end while

2.4.2.1 Setting $\hat{\boldsymbol{R}}_2$

Since $\hat{\boldsymbol{K}}_2$ is not assembled in practice, we cannot obtain $\hat{\boldsymbol{R}}_2$ via a direct null space computation of $\hat{\boldsymbol{K}}_2$. For a deterministic problem, rigid body modes of the physical domain span the null space of the stiffness matrix [82]. For three-dimensional problems, the maximum number of rigid body modes is six, while for two-dimensional deterministic problems this number reduces to three. Let $\boldsymbol{R}_2$ denote the matrix associated with the rigid body modes of the sub-domain $\mathcal{D}_2$ computed based on the geometry of $\mathcal{D}_2$, see, e.g. [183]. Then $\hat{\boldsymbol{R}}_2$ may be set as follows.

Proposition 2. *$\hat{\boldsymbol{R}}_2$ is an $r \times r$ block diagonal matrix whose (l,l) block, $l = 1, \cdots, r$, is given by*

$$\hat{\boldsymbol{R}}_2(l,l) = \boldsymbol{R}_2. \tag{2.25}$$

Proof. Since columns of $\boldsymbol{R}_2$ are linearly independent, columns of $\hat{\boldsymbol{R}}_2$ are also linearly independent as $\hat{\boldsymbol{R}}_2$ is block diagonal. On the other hand, as the columns of $\boldsymbol{R}_2$ form a basis for the null space of $\boldsymbol{K}_2(\boldsymbol{\xi})$, we have $\hat{\boldsymbol{K}}_2(l,l)\boldsymbol{R}_2 = \boldsymbol{0}$, $l = 1, \cdots, r$, where $\hat{\boldsymbol{K}}_2(l,l)$ is given in (2.13). Consequently $\hat{\boldsymbol{K}}_2\hat{\boldsymbol{R}}_2 = \boldsymbol{0}$, which, together with the linear independency of the columns of $\hat{\boldsymbol{R}}_2$, implies $\mathrm{range}(\hat{\boldsymbol{R}}_2) = \ker(\hat{\boldsymbol{K}}_2)$. □

2.4.3 Response statistics

In this section, the computation of response statistics based on the separated representation (2.9) is presented. Let us denote the mean and variance of the separated representation $\boldsymbol{u}_i^r = \sum_{l=1}^r \boldsymbol{u}_{0,i}^l\phi_1^l\phi_2^l$, $i = 1, 2$, by $\mathbb{E}[\boldsymbol{u}_i^r]$ and $\mathrm{var}[\boldsymbol{u}_i^r]$, respectively. For example, we can approximate the point-wise mean and the

second moment tensor of sub-domain solutions $\boldsymbol{u}_i$ by

$$\mathbb{E}[\boldsymbol{u}_i^r] = \sum_{l=1}^{r} \boldsymbol{u}_{0,i}^l \mathbb{E}_{\boldsymbol{\xi}_1}[\phi_1^l]\mathbb{E}_{\boldsymbol{\xi}_2}[\phi_2^l],$$
$$\mathbb{E}[\boldsymbol{u}_i^r \otimes \boldsymbol{u}_i^r] = \sum_{l=1}^{r}\sum_{l'=1}^{r} (\boldsymbol{u}_{0,i}^l \otimes \boldsymbol{u}_{0,i}^{l'})\mathbb{E}_{\boldsymbol{\xi}_1}[\phi_1^l\phi_1^{l'}]\mathbb{E}_{\boldsymbol{\xi}_2}[\phi_2^l\phi_2^{l'}], \tag{2.26}$$

respectively, where $\otimes$ denotes the Kronecker product of two vectors.

Alternatively, one may generate Monte Carlo estimates of the statistics of $\boldsymbol{u}_i^r$ by directly sampling the representation (2.9). We use the latter approach to compute the probability density function of the solution of interest in our numerical experiments.

2.4.4 Computational cost

We next elaborate on the computational complexity of one sweep of the ARR algorithm, which includes assembling and solving the deterministic update (2.12), as well as the stochastic updates (2.15) and (2.17). For the interest of simplicity, we consider the case of two coupled, non-floating sub-domains, each with d_1 independent random inputs. We further assume the sub-domain solutions are discretized such that they have the same number of spatial degrees of freedom $M_1 \gg 1$ and PC expansion order p. Following (1.5), this leads to $P_1 = (p+d_1)!/(p!d_1!)$ PC basis functions for the solution to (2.15) and (2.17). Moreover, we assume each $\boldsymbol{K}_i(\boldsymbol{\xi}_i)$, $i = 1, 2$, has at most ζ_K non-zero entries on each row, and is expanded in a PC basis of cardinality P_K.

For a rank r separated representation, the linear system (2.12) can be formed with complexity $\mathcal{O}\left(P_K(\zeta_S P_1 + \zeta_K M_1)r^2\right)$ and solved with cost $\mathcal{O}(M_1^{\beta_0} r^{\beta_0})$, where $\beta_0 \in [1, 3]$ depends on the choice of linear solver. Here, ζ_S is the maximum number of non-zeros on each row of sparse matrices $\boldsymbol{S}_k$ with entries $\boldsymbol{S}_k(i, j) = \mathbb{E}[\psi_k\psi_i\psi_j]$, where $i, j = 1, \ldots, P_1$ and $k = 1, \ldots, P_K$. The matrices $\boldsymbol{S}_k$ are generated once and utilized in computing the expectations in (2.13), and later in (2.16) and (2.18). The assembly and solution of the stochastic update (2.15), or (2.17), can be achieved with complexities $\mathcal{O}((P_K\zeta_S P_1 + P_K\zeta_K M_1 + \zeta_K M_1)r^2) \approx \mathcal{O}(P_K(\zeta_S P_1 + \zeta_K M_1)r^2)$ and $\mathcal{O}(P_1^{\beta_1} r^{\beta_1})$, respectively, where $\beta_1 \in [1, 3]$. We highlight that all the above complexity estimates depend on d_1, through P_1, and not the total dimensionality $d = 2d_1$.

As a comparison, the cost of forming and solving the linear system associated with order p Galerkin PC approximation of (2.5) may be as high as $\mathcal{O}(P_K\zeta_K M_1 P^2)$ and $\mathcal{O}(M_1^\beta P^\beta)$, respectively. Here, $P = (p+2d_1)!/(p!(2d_1)!)$ is given by (1.5) with $d = 2d_1$ and $\beta \in [1,3]$.

Notice that, for cases where d_1 is large (hence, $P_1 \ll P$) and the solution admits a low separation rank r, i.e., $r \ll P$, the cost of the proposed separated representation approach is significantly smaller than that of the standard Galerkin PC method. Table 2.1 summarizes the total computational costs of rank r separated representation – assuming ζ_r iterations are needed for a converged solution – as well as the standard Galerkin PC approximation of (2.5). The comparison of the actual cost of these two approaches as well as the existence of solutions with low separation ranks r are empirically shown in the numerical examples of next section.

Table 2.1: Computational cost of forming and solving linear systems arising in separated representation and Galerkin PC approximation of (2.5). Here, ζ_r denotes the number of iterations needed for a converged rank r separated representation. All the involved parameters are specified in Section 2.4.4.

Scheme	Forming	Solving
Separated representation	$\mathcal{O}\left(\zeta_r P_K(\zeta_S P_1 + \zeta_K M_1) r^2\right)$	$\mathcal{O}(\zeta_r(M_1^{\beta_0} r^{\beta_0} + P_1^{\beta_1} r^{\beta_1}))$
Standard Galerkin PC	$\mathcal{O}(P_K\zeta_K M_1 P^2)$	$\mathcal{O}(M_1^\beta P^\beta)$

2.5 Numerical Examples

In this section, we present two numerical examples for the verification of the proposed stochastic coupling framework. In the first example, we consider a linear elliptic PDE defined on a two-dimensional L-shaped (spatial) domain with random diffusion coefficient. The second example deals with the problem of deformation of a linear elastic cantilever beam with uncertain Young's modulus. In both cases, the uncertainly is represented by random fields taking statistically independent values on two and three non-overlapping sub-domains, respectively, in the first and second examples. To study the convergence of the separated representation on the entire domain, the mean and standard deviation (std) error measures

$$\epsilon_\mu^r = \frac{\left\|\mathbb{E}[\boldsymbol{u}_{ref}] - \mathbb{E}[\boldsymbol{u}^r]\right\|_{L_2(\mathbb{R}^M)}}{\left\|\mathbb{E}[\boldsymbol{u}_{ref}]\right\|_{L_2(\mathbb{R}^M)}} \quad \text{and} \quad \epsilon_\sigma^r = \frac{\left\|\mathrm{std}[\boldsymbol{u}_{ref}] - \mathrm{std}[\boldsymbol{u}^r]\right\|_{L_2(\mathbb{R}^M)}}{\left\|\mathrm{std}[\boldsymbol{u}_{ref}]\right\|_{L_2(\mathbb{R}^M)}} \tag{2.27}$$

are used. Here $\boldsymbol{u}^r$ is the rank r separated representation of the solution $\boldsymbol{u}$ to the coupled problem (2.5) and $\boldsymbol{u}_{ref}$ is the FE solution to the original problem (4.8). The reference solution $\boldsymbol{u}_{ref}$ is the PC expansion of $\boldsymbol{u}$ constructed with respect to all random inputs $\boldsymbol{\xi}$ and is obtained via Galerkin projection.

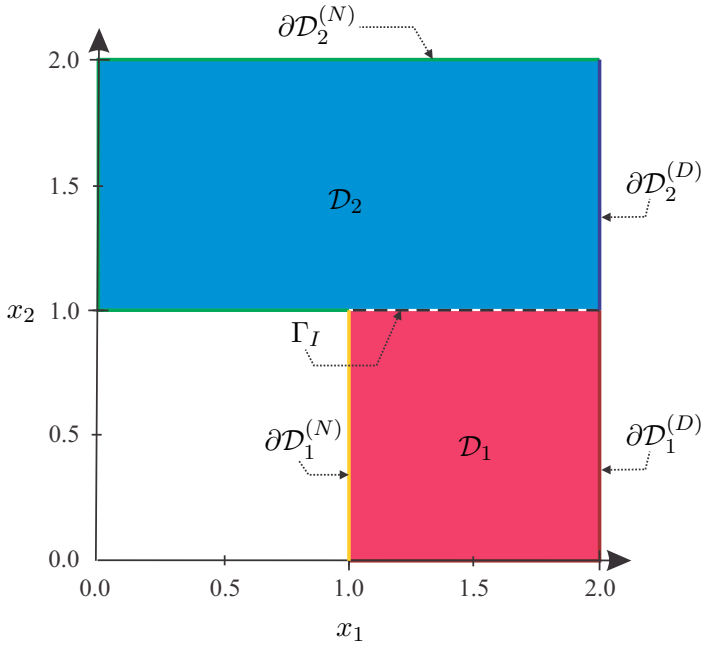


Figure 2.2: Geometry of the 2D elliptic problem over an L-shaped domain ($\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$ and $\mathcal{D}_1 \cap \mathcal{D}_2 = \emptyset$).

2.5.1 Example I: 2D elliptic PDE with random diffusion coefficient

Consider the following elliptic PDE with random diffusion coefficient κ,

$$\begin{aligned} -\nabla \cdot (\kappa(\boldsymbol{x}, \boldsymbol{\xi}) \nabla u(\boldsymbol{x}, \boldsymbol{\xi})) &= f(\boldsymbol{x}) && \boldsymbol{x} \in \mathcal{D}, \\ u(\boldsymbol{x}, \boldsymbol{\xi}) &= 0 && \boldsymbol{x} \in \partial\mathcal{D}^{(D)}, \\ \nabla u(\boldsymbol{x}, \boldsymbol{\xi}) \cdot \boldsymbol{n} &= 0 && \boldsymbol{x} \in \partial\mathcal{D}^{(N)}, \end{aligned} \tag{2.28}$$

defined over an L-shaped domain $\mathcal{D}$ depicted in Fig. 2.2. Dirichlet and Neumann boundary conditions are denoted by $\partial\mathcal{D}^{(D)}$ and $\partial\mathcal{D}^{(N)}$, respectively, and $\boldsymbol{n}$ is the unit normal vector to $\partial\mathcal{D}^{(N)}$.

We assume $\mathcal{D}$ is composed of two non-overlapping sub-domains $\mathcal{D}_1$ and $\mathcal{D}_2$, i.e., $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$ and

$\mathcal{D}_1 \cap \mathcal{D}_2 = \emptyset$ on which κ and f take different values. Specifically, we set

$$\log\left(\kappa(\boldsymbol{x},\boldsymbol{\xi}) - \kappa_0\right) = \begin{cases} G_1(\boldsymbol{x},\boldsymbol{\xi}_1) & \boldsymbol{x} \in \mathcal{D}_1 \\ G_2(\boldsymbol{x},\boldsymbol{\xi}_2) & \boldsymbol{x} \in \mathcal{D}_2 \end{cases} \tag{2.29}$$

and

$$f(\boldsymbol{x}) = \begin{cases} 10 & \boldsymbol{x} \in \mathcal{D}_1 \\ 0 & \boldsymbol{x} \in \mathcal{D}_2 \end{cases},$$

where G_1 and G_2 are statistically independent Gaussian random fields, and κ_0 is a small constant to ensure κ is bounded away from zero. Therefore, κ in (2.29) is a (shifted) lognormal random field on sub-domains $\mathcal{D}_1$ and $\mathcal{D}_2$. We represent G_1 and G_2 by the truncated Karhunen-Loève (KL) expansions

$$G_i(\boldsymbol{x},\boldsymbol{\xi}_i) = \bar{G}_i + \sum_{j=1}^{d_i} \sqrt{\tau_{i,j}}\, g_{i,j}(\boldsymbol{x})\xi_{i,j}, \qquad i = 1, 2,$$

where $\bar{G}_i$ is the mean of G_i, and $\{\xi_{1,j}\}_{j=1}^{d_1}$ and $\{\xi_{2,j}\}_{j=1}^{d_2}$ are i.i.d. standard Gaussian random variables. Additionally, $\{\tau_{i,j}\}_{j=1}^{d_i}$ and $\{g_{i,j}(\boldsymbol{x})\}_{j=1}^{d_i}$ are, respectively, d_i largest eigenvalues and the corresponding eigenfunctions of the Gaussian covariance kernel

$$C_i(\boldsymbol{x}_1,\boldsymbol{x}_2) = \sigma_i^2 \exp\left(-\frac{\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|_2^2}{l_i^2}\right), \qquad \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{D}_i. \tag{2.30}$$

Here, σ_i controls the variability of G_i and l_i is the correlation length of G_i. To exactly compute the expectations involving $\boldsymbol{K}_1(\boldsymbol{\xi}_1)$ and $\boldsymbol{K}_2(\boldsymbol{\xi}_2)$ in (2.13), (2.16), and (2.18), we decompose $\kappa(\boldsymbol{x},\boldsymbol{\xi})$, separately on each sub-domain, into Hermite polynomial chaos of order $2p_i$, i.e., $\kappa(\boldsymbol{x},\boldsymbol{\xi}_i) = \sum_{\boldsymbol{j}\in\mathscr{I}_{d_i,2p_i}} \kappa_{\boldsymbol{j},i}(\boldsymbol{x})\psi_{\boldsymbol{j}}(\boldsymbol{\xi}_i)$, $\boldsymbol{x} \in \mathcal{D}_i$. Here, p_i is the order of the PC expansion of the stochastic functions $\{\phi_i^l(\boldsymbol{\xi}_i)\}$ in (2.19). For the case of lognormal $\kappa(\boldsymbol{x},\boldsymbol{\xi}_i)$, the expansion coefficients $\{\kappa_{\boldsymbol{j},i}(\boldsymbol{x})\}_{\boldsymbol{j}\in\mathscr{I}_{d_i,2p_i}}$ are available analytically, see [246] or Appendix B.

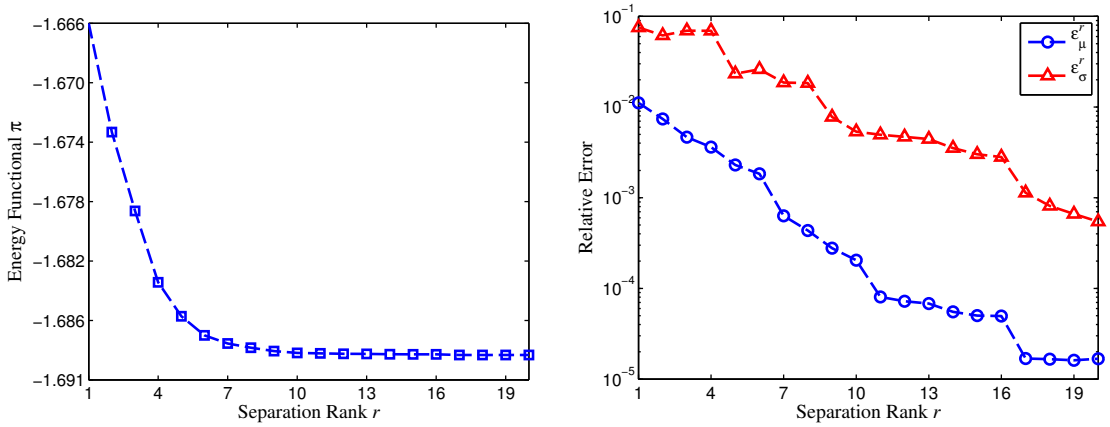
Table 2.2 summarizes the parameter values involved in the representation of κ.

Table 2.2: Assumed parameters for the PC representation of κ in (2.29).

d_1	d_2	p_1	p_2	$l_{c,1}$	$l_{c,2}$	$\bar{G}_1$	$\bar{G}_2$	σ_1	σ_2	κ_0
4	6	3	3	2/3	1/3	1	1	$1/2$	$1/2$	0.28

The FE discretization of (2.28) on each sub-domain is done using the FEniCS package [144]. In particular, we use linear triangle elements with a uniform mesh size $h_1 = h_2 = 1/20$ along x_1 and x_2.

To demonstrate convergence of the separated representation of the solution u to (2.28), we increase the separation rank r up to $r = 20$. Alternatively, we may prescribe a target accuracy ϵ for the residual norm ϵ^r_{res} in (2.11) and identify the corresponding separation rank r. In the latter approach, the solution refinement is achieved by decreasing ϵ.



(a) (b)

Figure 2.3: Energy functional π (a) and relative errors in mean and standard deviation (b) as a function of the separation rank r for the L-shaped problem. The errors are evaluated from (2.27).

Fig. 2.3(a) shows the values of the energy functional π with respect to increments in the separation rank r. While this is not generally the case, here an increase in r leads to a monotonic reduction in π. Fig. 2.3(b) presents the mean and standard deviation errors ϵ^r_μ and ϵ^r_σ, respectively, as defined in (2.27). As can be observed from this figure, there is not monotonicity in the reduction of mean and standard deviation errors as a function of r.

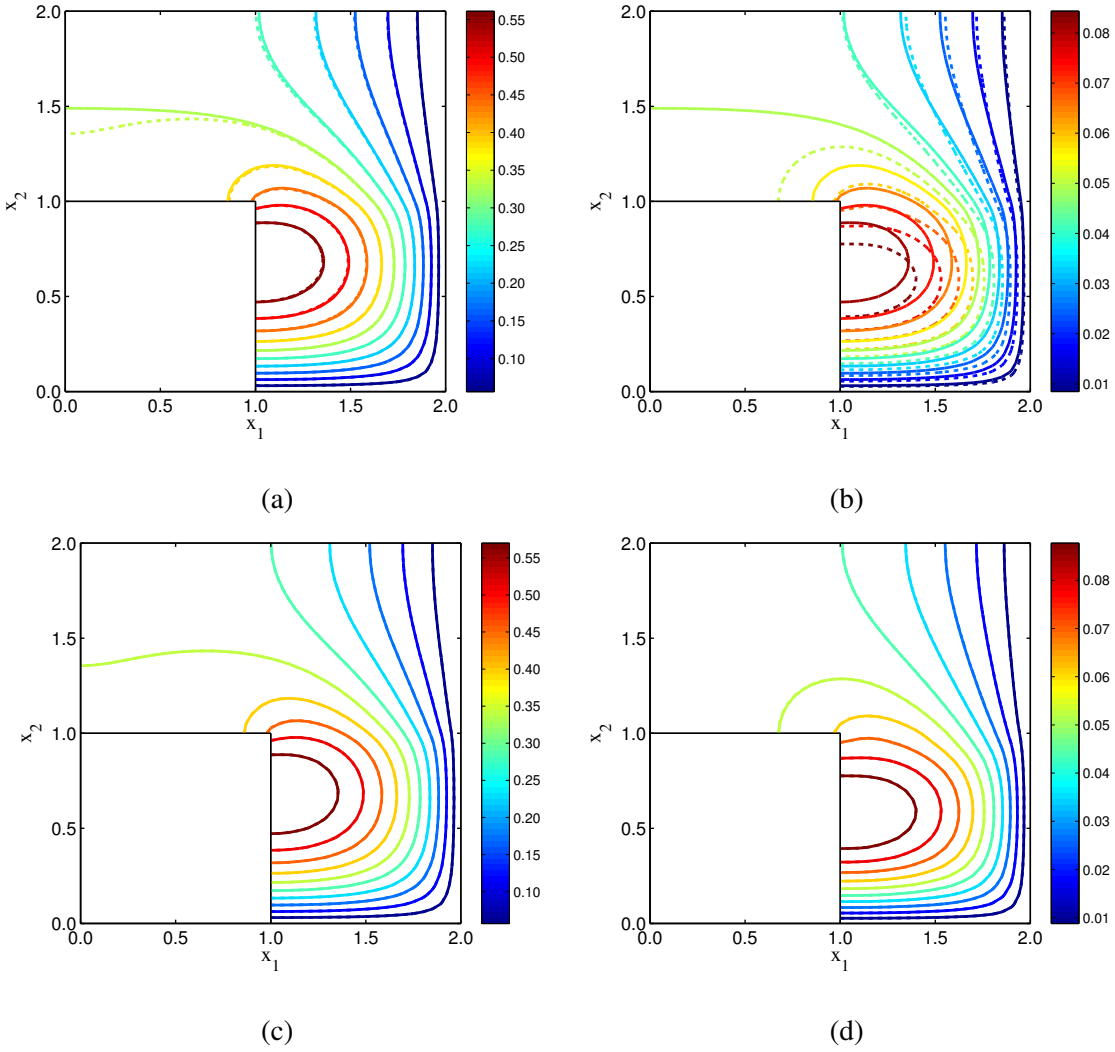


Figure 2.4: Contours of the solution mean and standard deviation obtained with separated representation (solid line) and the reference solution (dotted line) for the L-shaped problem. (a) Mean for $r = 1$; (b) Standard deviation for $r = 1$; (c) Mean for $r = 20$; (d) Standard deviation for $r = 20$.

In Figs. 2.4(a)-(b) we display the contours of the solution mean and standard deviation obtained from the separated representation with $r = 1$ as well as the reference solution. While approximation of the solution mean may be relatively accurate, the standard deviation has not yet converged with $r = 1$. By increasing the separation rank to $r = 20$, however, the approximation of these quantities improves considerably as can be observed from Figs. 2.4(c)-(d).

In Fig. 2.5, the probability density function (PDF) of the separated representation of solution at

$(x_1, x_2) = (1.0, 0.5)$ is compared to that of the reference solution. For the case of small separation rank, $r = 1$, there is a notable difference between the two PDFs. However, this disagreement reduces considerably as the separation rank is increased to $r = 20$.

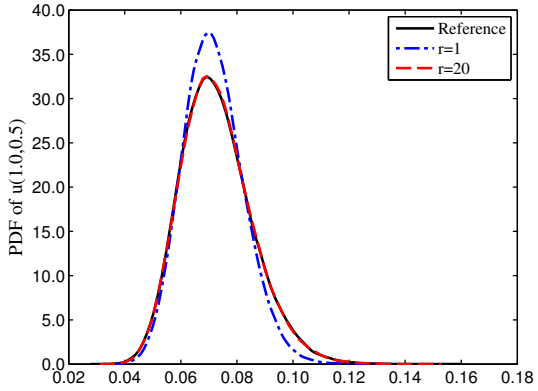


Figure 2.5: PDF of $u(1.0, 0.5)$ for the L-shaped problem. A comparison between the rank r separated representation and the reference solution.

As discussed in Section 2.4, the stochastic functions $\{\phi_1^l\}_{l=1}^r$ and $\{\phi_2^l\}_{l=1}^r$ are not fixed *a priori* and are computed by iteratively solving for the saddle point of π. Therefore, their representation is problem-dependent. Fig. 2.6 presents the PDFs of ϕ_1^r and ϕ_2^r when $r = 1, 3, 7, 12$. As it can be observed, these stochastic functions differ from each other and are not identically distributed. The actual stochastic basis functions in (2.9) are products of these functions, i.e., $\phi_1^r \phi_2^r$. Fig. 2.7 shows the PDFs of a number of these basis functions. We note that these basis functions are generally different from the PC basis functions.

In order to compare the computational cost of the separated representation approach with the standard PC expansion via SG, we report the ratio of their execution times in Fig. 2.8. By increasing r, the sizes of the stochastic and deterministic update systems and, consequently, their solution times increases. However, as the number of alternations for each separation rank r may not depend on r, the run time of ARR is not monotonic with respect to r. This may be observed from Fig. 2.8. The run time of obtaining the PC reference solution was about 1384 minutes, which was three orders of magnitudes larger than that of these ARR approach, see Fig. 2.8. We note that here both solutions were obtained via serial implementations.

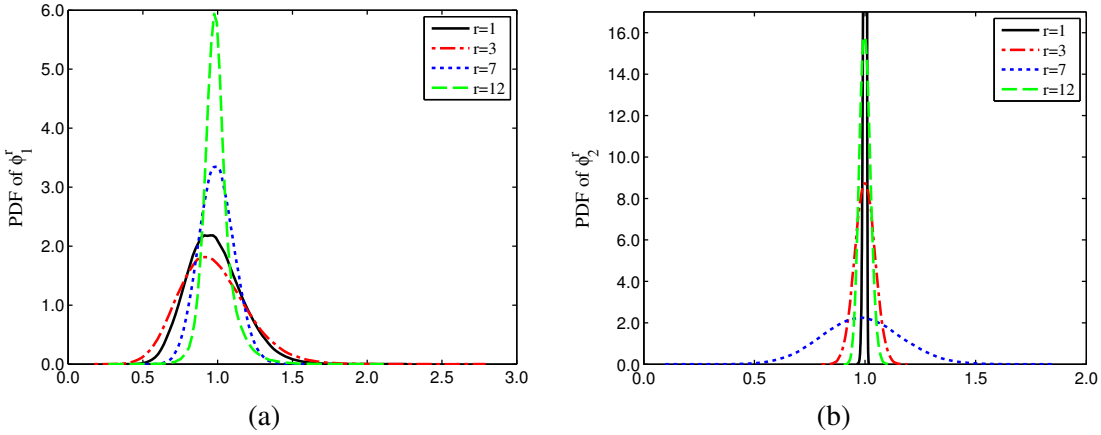


Figure 2.6: PDFs of the normalized ϕ_1^r and ϕ_2^r in the separated representation of the solution to the L-shaped problem. (a) PDF of the normalized ϕ_1^r, $r = 1, 3, 7, 12$; (b) PDF of the normalized ϕ_2^r, $r = 1, 3, 7, 12$. Each ϕ_i^r is normalized such that it has unit second moment. The PDF of ϕ_2^1 is not fully shown here.

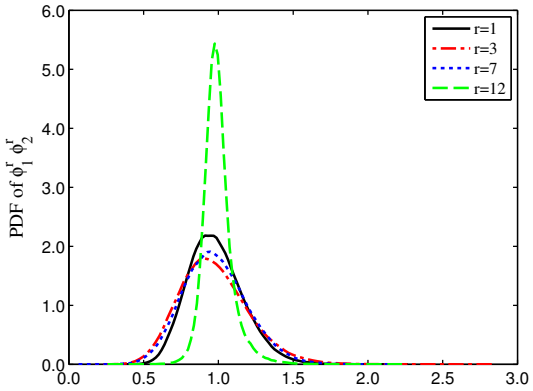


Figure 2.7: PDFs of the normalized stochastic basis functions $\phi_1^r \phi_2^r$, $r = 1, 3, 7, 12$, for the L-shaped problem. Here each basis function is normalized to have unit second moment.

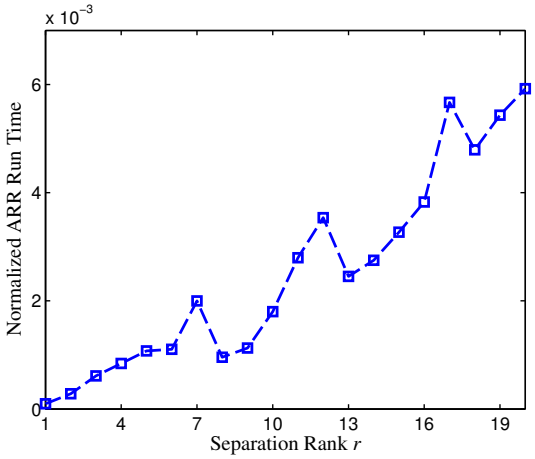


Figure 2.8: Execution time of the ARR algorithm as a function of the separation rank r for the L-shaped problem. The normalization is with respect to the run time of the Galerkin-based PC method.

2.5.2 Example II: 2D stochastic linear elasticity

For the second numerical example, we consider the linear elasticity problem,

$$\begin{aligned} -\nabla \cdot \boldsymbol{\sigma}\left(\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{\xi})\right) &= \mathbf{0} && \boldsymbol{x} \in \mathcal{D}, \\ \boldsymbol{u}(\boldsymbol{x}, \boldsymbol{\xi}) &= \mathbf{0} && \boldsymbol{x} \in \partial\mathcal{D}^{(D)}, \\ \boldsymbol{\sigma}\left(\boldsymbol{u}(\boldsymbol{x}, \boldsymbol{\xi})\right) \boldsymbol{n} &= \boldsymbol{t}(\boldsymbol{x}) && \boldsymbol{x} \in \partial\mathcal{D}^{(N)}, \end{aligned} \tag{2.31}$$

describing the deformation of the cantilever beam shown in Fig. 2.9. Here, $\boldsymbol{u}$ denotes the displacement vector field, $\boldsymbol{\sigma}$ is the stress tensor, $\boldsymbol{n}$ is the unit normal vector to the boundary, and $\boldsymbol{t}$ is the traction vector. We assume that $\boldsymbol{t} = (0, -0.1)$ on the edge corresponding to $x_2 = 1.0$ and $\boldsymbol{t} = (0, 0)$ elsewhere. The stress tensor $\boldsymbol{\sigma}$ is related to the strain tensor $\boldsymbol{e} = \left(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^T\right)/2$ via the isotropic linear elastic stress-strain relation

$$\boldsymbol{\sigma} = \frac{E}{1+\nu}\left(\boldsymbol{e} + \frac{\nu}{1-2\nu}\,\mathrm{tr}(\boldsymbol{e})\boldsymbol{I}\right),$$

where $\nu = 0.3$ is the Poisson's ratio. Here, $\mathrm{tr}(\cdot)$ denotes the trace operator of a tensor and $\boldsymbol{I}$ is the identity tensor of order two. The Young's modulus $E(\boldsymbol{x}, \boldsymbol{\xi})$ is the source of uncertainty in (2.31) and is assumed to take statistically independent values over sub-domains $\mathcal{D}_1$, $\mathcal{D}_2$ and $\mathcal{D}_3$ shown in Fig. 2.9. Specifically,

$$E(\boldsymbol{x}, \boldsymbol{\xi}) = \begin{cases} E_1(\boldsymbol{x}, \boldsymbol{\xi}_1) & \boldsymbol{x} \in \mathcal{D}_1 \\ E_2(\boldsymbol{x}, \boldsymbol{\xi}_2) & \boldsymbol{x} \in \mathcal{D}_2 \\ E_3(\boldsymbol{x}, \boldsymbol{\xi}_3) & \boldsymbol{x} \in \mathcal{D}_3 \end{cases}, \tag{2.32}$$

where each $E_i(\boldsymbol{x}, \boldsymbol{\xi}_i)$, $i = 1, 2, 3$, is represented by the series

$$E_i(\boldsymbol{x}, \boldsymbol{\xi}_i) = \bar{E}_i + \sum_{j=1}^{d_i} \sqrt{\tau_{i,j}}\, g_{i,j}(\boldsymbol{x}) \xi_{i,j}, \qquad \boldsymbol{x} \in \mathcal{D}_i. \tag{2.33}$$

Here, $\{\tau_{i,j}\}_{j=1}^{d_i}$ and $\{g_{i,j}(\boldsymbol{x})\}_{j=1}^{d_i}$ are d_i largest eigenvalues and the corresponding eigenfunctions of the Gaussian covariance kernel given in (2.30). We assume that $\{\xi_{i,j}\}_{j=1}^{d_i}$, $i = 1, 2, 3$, are i.i.d. uniform random variables $U(-1, 1)$. The list of parameters used in the analyses of this example is given in Table 2.3. These choices ensure that all realizations of E are strictly positive on $\mathcal{D}$.

Similar to the L-shaped problem, FE discretizations are done via the FEniCS project [144] and by using linear triangle elements with a uniform grid size $h_1 = h_2 = 1/10$ along x_1 and x_2 directions.

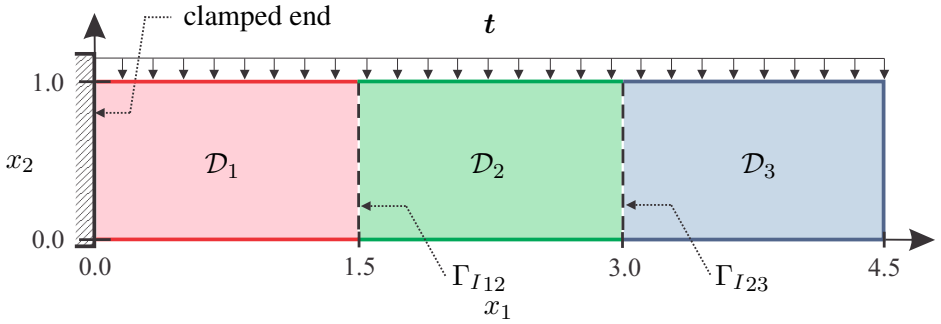


Figure 2.9: Geometry of the 2D cantilever beam composed of three non-overlapping sub-domains $\mathcal{D}_1$, $\mathcal{D}_2$ and $\mathcal{D}_3$. The two interfaces are: $\Gamma_{I12} = \partial\mathcal{D}_1 \cap \partial\mathcal{D}_2$ and $\Gamma_{I23} = \partial\mathcal{D}_2 \cap \partial\mathcal{D}_3$. Young's moduli of the sub-domain materials are statistically independent random fields.

Table 2.3: Assumed parameters for the description of Young's modulus E in (2.32) and (2.33).

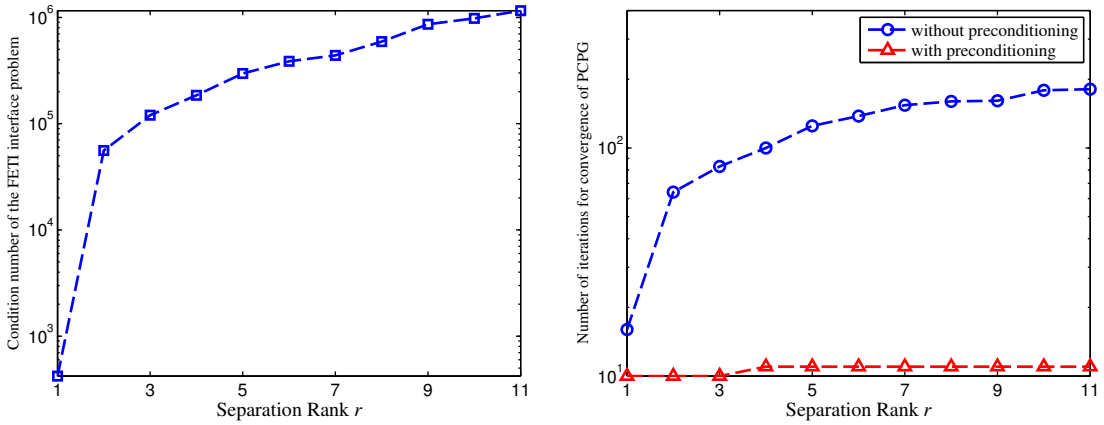
d_1	d_2	d_3	$l_{c,1}$	$l_{c,2}$	$l_{c,3}$	$\bar{E}_1$	$\bar{E}_2$	$\bar{E}_3$	σ_1	σ_2	σ_3
5	8	7	2/3	1/2	1/3	100	100	100	35	35	35

Legendre PC expansions of $\{\phi_i^l\}_{l=1}^r$, $i = 1, 2, 3$, with degree $p = 3$ were found sufficient for the solution of update equations (2.15) and (2.17).

We note that in the present test case, $\mathcal{D}_2$ and $\mathcal{D}_3$ are floating sub-domains with no Dirichlet boundary conditions; therefore, the PCPG solver described in Section 2.4.2 is applied to compute the vectors of Lagrange multipliers $\{\boldsymbol{\lambda}_0^l\}_{l=1}^r$ in (2.12). We refer to Algorithm 2 for the implementation details of the PCPG solver. In our computations, we have considered $\epsilon_{PCPG} = 10^{-6}$ as the stopping criterion for the PCPG solver. The condition number of $\hat{\boldsymbol{F}}_I$ determines the number of PCPG iterations and generally increases as a function of the separation rank r, hence asking for a preconditioner. Selection of an effective preconditioner $\bar{\boldsymbol{F}}_I^{-1}$, however, is generally a non-trivial task in the FETI approach. Among several existing choices, e.g., [197, 133, 244, 193, 92, 38], we found the preconditioner of [133],

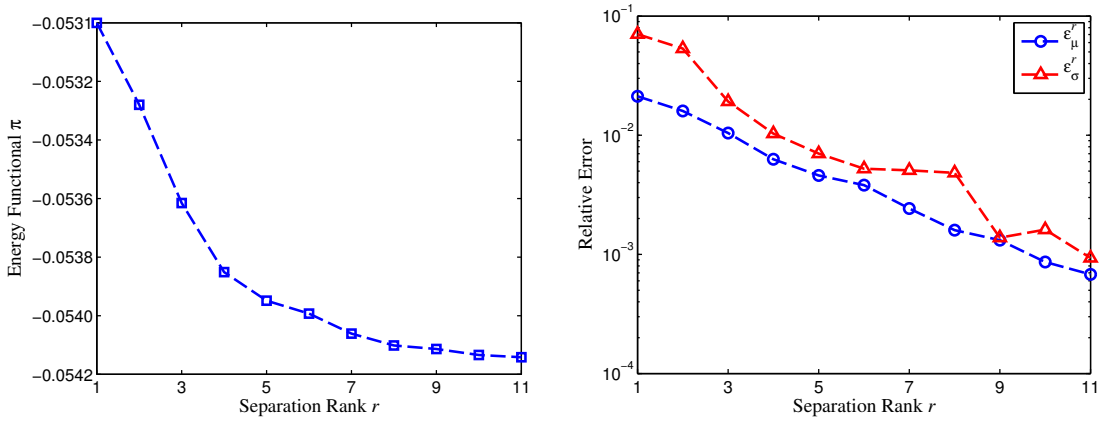
$$\bar{\boldsymbol{F}}_I^{-1} = \left(\sum_{i=1}^{3} \hat{\boldsymbol{C}}_i^T \hat{\boldsymbol{C}}_i\right)^{-1} \left(\sum_{i=1}^{3} \hat{\boldsymbol{C}}_i^T \hat{\boldsymbol{K}}_i \hat{\boldsymbol{C}}_i\right) \left(\sum_{i=1}^{3} \hat{\boldsymbol{C}}_i^T \hat{\boldsymbol{C}}_i\right)^{-1}, \tag{2.34}$$

particularly effective for our purpose.



(a) (b)

Figure 2.10: Convergence properties of the PCPG algorithm used for the beam problem. (a) Condition number of $\hat{\boldsymbol{F}}_I$ vs. the separation rank; (b) Number of the iterations required for the PCPG algorithm to converge with and without preconditioning vs. the separation rank.



(a) (b)

Figure 2.11: Energy functional π (a) and relative errors in mean and standard deviation of the displacement (b) as a function of the separation rank r for the beam problem. The errors are evaluated from (2.27).

In Fig. 2.10(a), we show the dependence of the condition number of $\hat{\boldsymbol{F}}_I$ on r for the present example. A rapid increase in the condition number of $\hat{\boldsymbol{F}}_I$ can be observed when r is increased. Fig. 2.10(b) displays the number of PCPG iterations required to reach $\epsilon_{PCPG} = 10^{-6}$ with the preconditioner (2.34) and when no preconditioner is used, i.e., $\bar{\boldsymbol{F}}_I^{-1} = \boldsymbol{I}$. We observe that the choice of preconditioner (2.34) makes

the convergence of the PCPG solver almost independent of r. This is particularly crucial when one is dealing with problems in which r is large. However, we note that further analysis is needed to confirm the effectiveness of (2.34).

Fig. 2.11 illustrates the convergence of the separated representation. Similar to the L-shaped problem, increasing the separation rank r results in a monotonic descries of the energy functional π, see Fig. 2.11(a). The relative mean and standard deviation errors ϵ_μ^r and ϵ_σ^r of the displacement are computed from (2.11) and are plotted against r in Fig. 2.11(b).

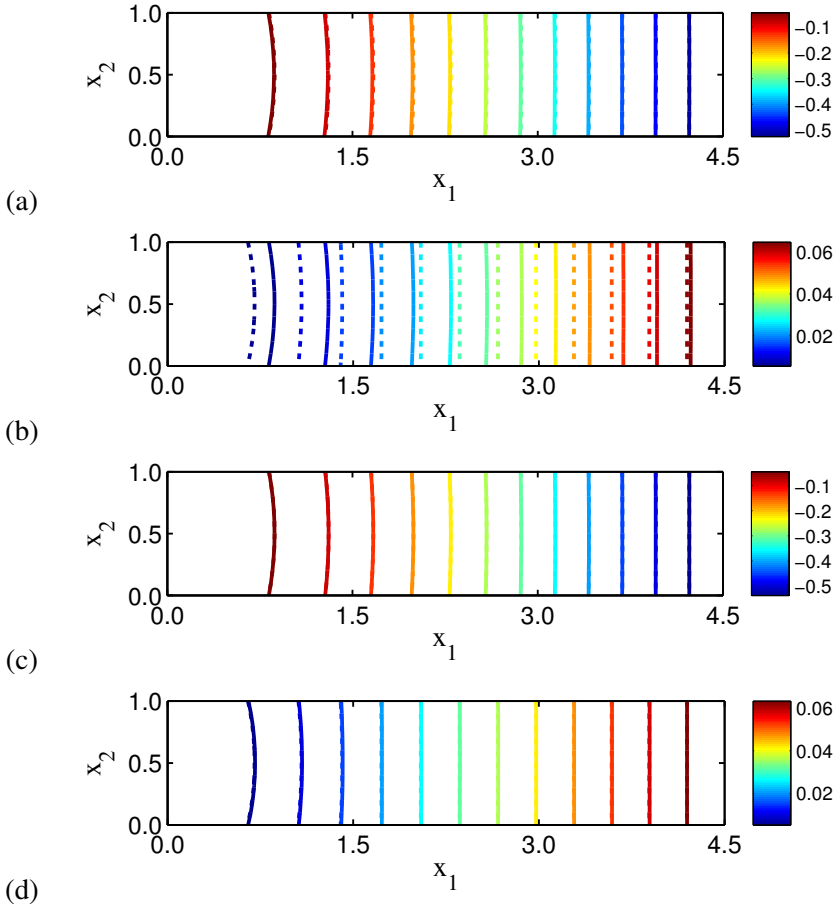


Figure 2.12: Contours of the mean and standard deviation of vertical displacement obtained with separated representation (solid line) and the reference solution (dotted line) for the beam problem. (a) Mean for $r = 1$; (b) Standard deviation for $r = 1$; (c) Mean for $r = 11$; (d) Standard deviation for $r = 11$.

In Fig. 2.12, we display the contours of the mean and standard deviation of the vertical displacement obtained from the separated representation, and compare them with those of the reference solution. The

mean is captured fairly accurately with a rank one approximation. While for $r = 1$ a considerable deviation from the standard deviation of the reference solution is observed, the rank $r = 11$ approximation agrees well with the reference solution (Fig. 2.12(b) and Fig. 2.12(d)). Despite the high-dimensionality of the solution, i.e., $d = 20$, we note that we only require a low separation rank, $r = 11$, to accurately approximate the solution mean and standard deviation.

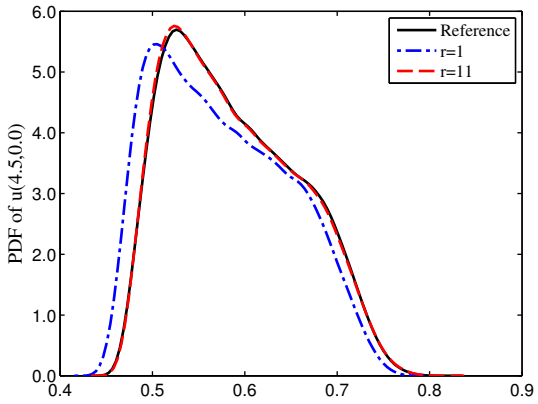


Figure 2.13: PDF of the total displacement at $(x_1, x_2) = (4.5, 0.0)$. A comparison between the separated approximation and the reference solution.

Fig. 2.13 compares the PDFs of the separated approximation of the total displacement at $(x_1, x_2) = (4.5, 0.0)$, when $r = 1$ and $r = 11$, to the PDF of the corresponding reference solution. An almost identical agreement between these PDFs is observed when $r = 11$.

To demonstrate that the stochastic functions ϕ_1^r, ϕ_2^r, and ϕ_3^r in (2.9) depend on the problem at hand, we present the PDFs of these quantities, for $r = 1, 3, 8, 11$, in Fig. 2.14. As it can be seen, these PDFs are different from those of the L-shaped problem. Additionally, Fig. 2.15 reports the PDF of the stochastic basis functions $\phi_1^r \phi_2^r \phi_3^r$ for the above r values.

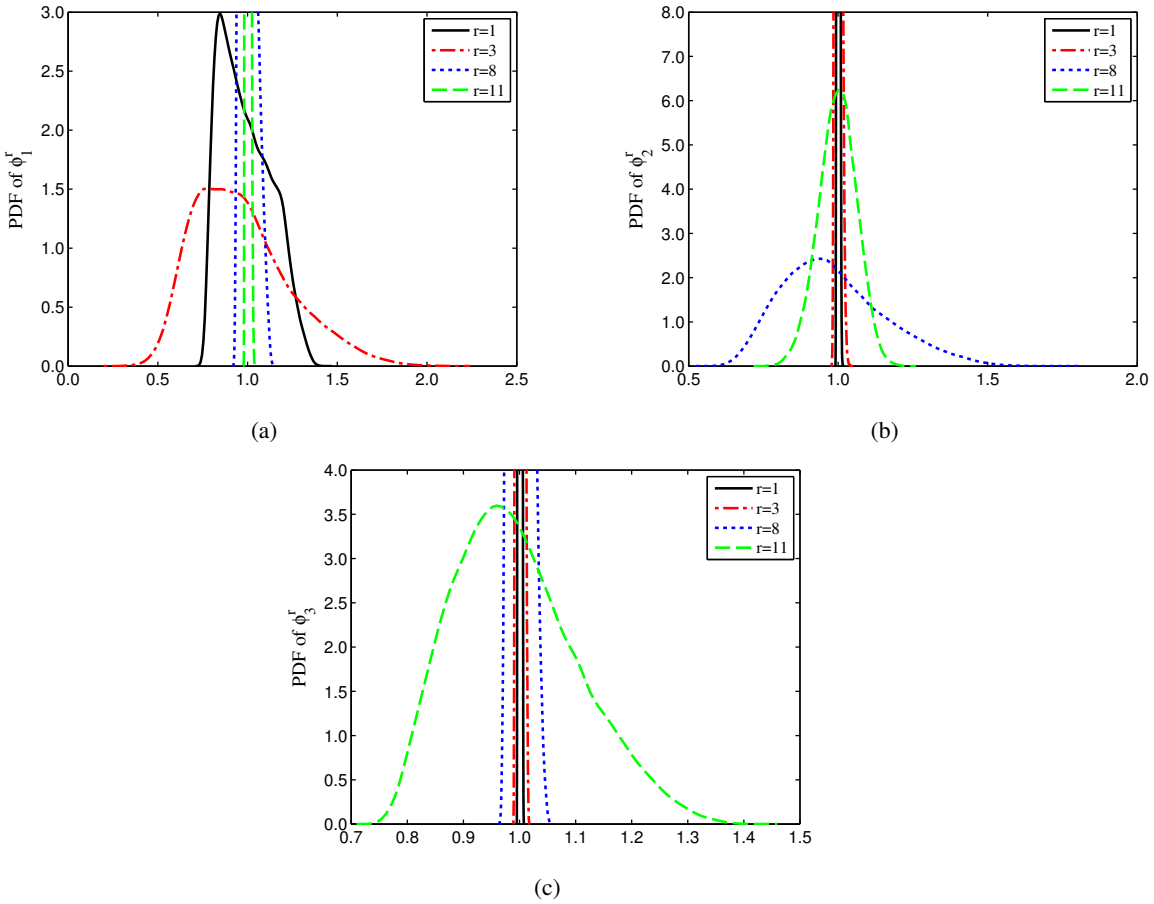


Figure 2.14: PDFs of the normalized stochastic functions ϕ_1^r, ϕ_2^r and ϕ_3^r, $r = 1, 3, 8, 11$, for the beam problem. The normalized stochastic functions have unit second moment. (a) PDFs of the normalized ϕ_1^r; (b) PDFs of the normalized ϕ_2^r; (c) PDFs of the normalized ϕ_3^r. Some PDFs are not shown completely.

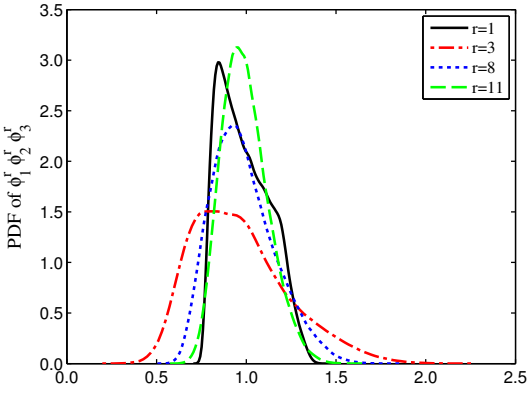


Figure 2.15: PDFs of the normalized stochastic basis functions $\phi_1^r \phi_2^r \phi_3^r$, $r = 1, 3, 8, 11$, for the beam problem. Each $\phi_1^r \phi_2^r \phi_3^r$ has unit second moment.

Fig. 2.16 shows the execution time of the ARR algorithm, for multiple r values, normalized by the execution time of of the SG approach for the beam solution. The serial implementation of the SG approach took about 628 minutes. As may be observed from Fig. 2.16, the separated representation approach leads to two orders of magnitude reduction in the computation time as compared to the PC counterpart.

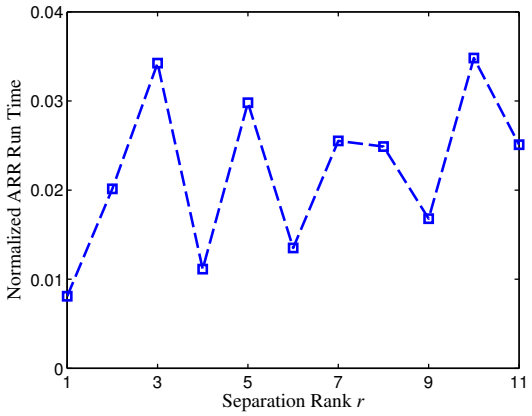


Figure 2.16: Execution time of the ARR algorithm as a function of the separation rank r for the beam problem. The normalization is with respect to the run time of the Galerkin-based PC method.

2.6 Summary and Conclusions

In this chapter, we developed a computational framework for the propagation of uncertainty through coupled domain problems. The proposed approach hinges on the construction of a solution-adaptive stochastic basis that is of separated form with respect to the random inputs characterizing the uncertainty in each sub-domain. Such a separated construction of stochastic basis is achieved through a sequence of approximations with respect to the dimensionality, i.e., number of random inputs, of each individual sub-domain and not the combined dimensionality. This leads to a *partitioned* treatment of the stochastic space and, consequently, a higher scalability of the method as compared with standard uncertainty propagation approaches, such as those based on direct polynomial chaos expansions. For situations where the cardinality of the separated basis – here referred to as the separation rank – is small, the proposed approach provides a reduced order representation of the coupled field solution.

The deterministic coefficients associated with each separated stochastic basis capture the spatial variability of the solution and are computed via the standard FETI approach. Therefore, the method achieves a high level of parallelism while requiring no intrusion in each sub-domain solver. Although our present formulation of domain coupling is based on the standard FETI approach, we foresee no major technical difficulties in employing more advanced domain coupling schemes.

The performance of the proposed framework was explored through its application to two linear elliptic PDEs with high-dimensional random inputs. Both problems were defined on physical domains consisting of two coupled sub-domains with independent sources of uncertainty. In both cases, despite the high-dimensionality of the random inputs, accurate estimations of the solution statistics were achieved with relatively low separation ranks, thus demonstrating the effectiveness of the present approach.

The stochastic expansion of this study based on the separated representations may also be applied to other coupled problems, such as fluid structure interaction (FSI), involving uncertainty. However, different numerical strategies for the construction of the separated basis may be required.

Chapter 3

Uncertainty Quantification of Lithium-ion Batteries

3.1 Introduction

Up to date, the majority of the LIB physics-based simulations have treated the underlying model parameters deterministically and ignored the effects of uncertainties in the model parameters on the performance of LIBs. There are a few works in the literature which have addressed the variability of LIB physics-based model parameters. In [207], PC expansion is employed within the Newman's model to demonstrate the reduction in the cell potentials as a result of the uncertainty in the particle size of the anode electrode. Effects of variability in cycling rate, particle size, diffusivity, and electrical conductivity of the cathode electrode on the LIB performance was examined in [65], where surrogate models are developed using the Newman's model together with techniques such as Kriging, polynomial response, and radial-basis neural networks. It was found that the randomness in electrical conductivity has the minimal effects on the cell performance, while the impact of the remaining parameters depend on the the cycling rate. In [208], impedance spectroscopy [222] is utilized to compare the relative importance of randomness in porosity, particle size, length and tortuosity of the cathode and separator using Nyquist stability criterion. It was shown that the particle size has the largest impact on the fluctuations in the impedance. In [192], a reduced order LIB model presented in [236] and a Bayesian inference technique are used to estimate the effective kinetic and transport parameters from experimental data. These works either employ a reduced order model of the LIB or only consider uncertainties in a few number of model parameters to avoid high computational costs.

This chapter presents a sampling-based PC approach to study the effects of uncertainty in various model parameters on the cell capacity, voltage, and concentrations of an LIB described by the Newman's

model [171, 62, 63]. The proposed PC approach, first introduced in [59], relies on the sparsity of expansion coefficients to accurately compute the statistics of quantities of interest with a small number of battery simulations. Through its application to an $LiC_6/LiCoO_2$ LIB model, we demonstrate that, unlike the previously mentioned works on UQ of LIBs, the proposed PC approach is capable of taking into account a large number of uncertain parameters. While the proposed PC-based UQ framework is general, the results we present are specific to the particular $LiC_6/LiCoO_2$ cell we consider and the the choices of uncertainty models for its parameters. The latter are, as much as possible, identified from the reported experiments on identical or similar cells/materials.

Additionally, this UQ framework enables performing a global sensitivity analysis (SA) to identify the most important uncertain parameters affecting the variability of the output quantities. Such an analysis may be used toward reducing cell-to-cell variations and designing more efficient and targeted quality control procedures to reduce the manufacturing cost of LIBs [208, 218]. We also review the standard experimental techniques used for measuring the model parameters and discuss their impacts on the cell capacity and/or power.

It is worth highlighting that the present work is focused on modeling and propagation of parametric uncertainties, as opposed to model form uncertainties which may consider multiple competing models.

3.2 LIB Governing Equations

An LIB schematic is presented in Fig. 3.1. During the discharge process, Lithium ions in the solid particles of the anode diffuse to the particle surface where they oxidize into Li^+ ions and electrons and transfer to the electrolyte liquid (deintercalation). Electrons flow through the external circuit to the positive electrode. Meanwhile, Li^+ ions travel via diffusion and migration through the electrolyte and separator to the cathode where they are reduced and diffused into the solid particles (intercalation). When the LIB is charged, this process is reversed.

The Newman's LIB model is developed based on porous electrode and concentrated solution theories [171, 62, 63]. The governing equations of this model for the salt concentration in liquid phase c, lithium concentration in solid phase c_s, liquid phase potential ϕ_e and solid phase potential ϕ_s are presented in

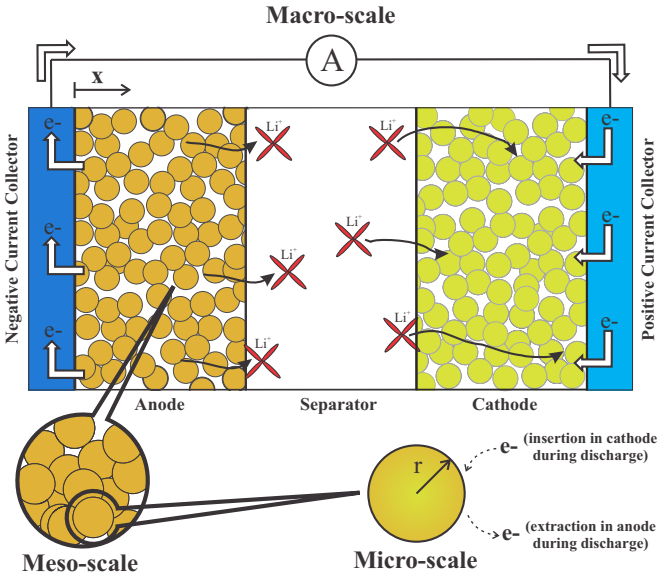


Figure 3.1: Schematic of a full cell LIB.

Table 3.1. Symbols used in this table are defined in the nomenclature at the end of this chapter. The assumptions made in this model are summarized as: (i) transport properties are independent of temperature and simulations are isothermal; (ii) D, D_s and t_+^0 do not depend on the concentrations; (iii) volume changes within the cell are ignored; (iv) zero surface electrolyte inter-phase (SEI) resistance is considered; (v) no double layer capacitive effects are considered; and (vi) solid particles are spherical.

In the literature, numerical techniques such as finite difference [62], finite volume [186], and finite elements [249] have been used to solve this system of coupled non-linear equations simultaneously. In order to reduce the computational cost of the LIB simulations, using the concepts of particular and homogeneous solutions to Ordinary Differential Equations (ODEs), Reimers [194] suggested a decoupled formulation of Newman's model, which we use in this study. For the sake of completeness, this decoupling technique is presented in the following.

The decoupling procedure starts with combining the potential equations for the solid phase (3.4), liquid phase (3.3) and the Butler-Volmer (BV) equation (3.5) in order to first reduce the number of equations to be solved. This is motivated by the fact that in each electrode, there are only three tightly coupled quantities, i.e., c, c_s and the over-potential $\eta = \phi_s - \phi_e - V$, which need to be solved simultaneously [194].

Table 3.1: Coupled non-linear governing equations of LIB.

	Governing equation		Boundary conditions
Electrolyte phase diffusion	$\frac{\partial(\epsilon c)}{\partial t} = \nabla(\epsilon D^{\text{eff}} \nabla c) + \frac{1-t_+^0}{F} j_{vol}$	(3.1)	$\nabla c\|_{x=0} = \nabla c\|_{x=L} = 0$
Solid phase diffusion	$\frac{\partial c_s}{\partial t} = \frac{1}{r^2}\frac{\partial}{\partial r}\Big(D_s r^2 \frac{\partial}{\partial r} c_s\Big)$	(3.2)	$\nabla c_s\|_{r=0} = 0$ $\nabla c_s\|_{r=r_s} = -\frac{j_{vol}}{aFD_s}$
Liquid phase potential	$\nabla(\kappa^{\text{eff}} \nabla \phi_e) - \nabla(\kappa_D^{\text{eff}} \nabla \ln c) + j_{vol} = 0$	(3.3)	$\nabla \phi_e\|_{x=0} = \nabla \phi_e\|_{x=L} = 0$ $\phi_e\|_{x=L} = 0$
Solid phase potential	$\nabla(\sigma^{\text{eff}} \nabla \phi_s) - j_{vol} = 0$	(3.4)	$\nabla \phi_s\|_{x=0} = \nabla \phi_s\|_{x=L} = \frac{-I}{\sigma^{\text{eff}}}$ $\nabla \phi_s\|_{x=L_a} = \nabla \phi_s\|_{x=L_a+L_s} = 0$
Reaction kinetics	$j_{vol} = a i_{ex}\Big[\exp\Big(\frac{0.5F\eta}{RT}\Big) - \exp\Big(-\frac{0.5F\eta}{RT}\Big)\Big]$ $i_{ex} = Fk(c_s^{\text{surf}})^{0.5}(c_{s,max} - c_s^{\text{surf}})^{0.5}(c)^{0.5}$	(3.5)	

Consequently, the following non-linear equation for the over-potential η can be obtained

$$\begin{aligned}\nabla(\sigma^h \nabla \eta) = a i_{ex}\Big[\exp\Big(\frac{0.5F\eta}{RT}\Big) - \exp\Big(-\frac{0.5F\eta}{RT}\Big)\Big] \\ -\nabla\Big[\frac{I}{\sigma^{\text{eff}}} + \frac{\kappa_D^{\text{eff}}}{\kappa^{\text{eff}}} \nabla \ln c(\eta) + \nabla V(\eta)\Big],\end{aligned} \tag{3.6}$$

subjected to the boundary conditions

$$\nabla \eta|_{x=0,L} = -\Big[\frac{I}{\sigma^{\text{eff}}} + \nabla V(\eta)\Big]_{x=0,L}, \tag{3.7}$$

$$\nabla \eta|_{x=L_a, L_a+L_s} = \Big[\frac{I}{\kappa^{\text{eff}}} - \frac{\kappa_D^{\text{eff}}}{\kappa^{\text{eff}}} \nabla \ln c(\eta) - \nabla V(\eta)\Big]_{x=L_a, L_a+L_s}. \tag{3.8}$$

Here, σ^h is the harmonic mean conductivity defined by $\frac{1}{\sigma^h} = \frac{1}{\sigma^{\text{eff}}} + \frac{1}{\kappa^{\text{eff}}}$. So far, instead of solving the coupled system of equations in Table 3.1, one needs to solve a system of equations including Eqs. (3.1), (3.2) and (3.6) simultaneously. We note that all of the nonlinearities are isolated in Eq. (3.6) for the over-potential.

The next step of deriving the reformulation is decoupling these three equations via the concept of particular and homogeneous solutions to ODEs. Application of the implicit finite time differencing to the

solid phase diffusion equation (3.2) results in the following non-homogeneous ODE with non-homogeneous BCs

$$c_s(r,t+\Delta t) - \frac{\Delta t}{r^2}\frac{\partial}{\partial r}\Big(D_s r^2 \frac{\partial}{\partial r} c_s(r,t+\Delta t)\Big) = c_s(r,t), \tag{3.9}$$

$$\nabla c_s|_{r=0} = 0, \qquad \nabla c_s|_{r=r_s} = -\frac{j_{vol}}{aFD_s}. \tag{3.10}$$

One may write the general solution to Eq. (3.9) as

$$c_s(r,t+\Delta t) = c_s^o(r,t+\Delta t) + \frac{j_{vol}}{aF} c_s^j(r,t+\Delta t), \tag{3.11}$$

where c_s^o is the solution to the following non-homogeneous ODE with homogeneous BCs

$$c_s^o(r,t+\Delta t) - \frac{\Delta t}{r^2}\frac{\partial}{\partial r}\Big(D_s r^2 \frac{\partial}{\partial r} c_s^o(r,t+\Delta t)\Big) = c_s(r,t), \tag{3.12}$$

$$\nabla c_s^o|_{r=0} = 0, \qquad \nabla c_s^o|_{r=r_s} = 0, \tag{3.13}$$

and c_s^j is the solution to the following the homogeneous ODE with non-homogeneous BCs

$$c_s^j(r,t+\Delta t) - \frac{\Delta t}{r^2}\frac{\partial}{\partial r}\Big(D_s r^2 \frac{\partial}{\partial r} c_s^j(r,t+\Delta t)\Big) = 0, \tag{3.14}$$

$$\nabla c_s^j|_{r=0} = 0, \qquad \nabla c_s^j|_{r=r_s} = -\frac{1}{D_s}. \tag{3.15}$$

A similar approach may be employed for decoupling the solution of the electrolyte phase diffusion equation (3.1) from the over-potential solution η and the volumetric pore wall flux j_{vol}. The only difference here is that j_{vol} in Eq. (3.1) is a position dependent source term while in Eq. (3.2) it appears on the boundary conditions. Consequently, the general solution to the discretized version of Eq. (3.1) in the time domain can be obtained via

$$c(x,t+\Delta t) = c^o(x,t+\Delta t) + \frac{1-t_+}{F}\int j_{vol}(x_0) c^j(x,x_0,t+\Delta t) dx_0. \tag{3.16}$$

Here in Eq. (3.16), c^o and c^j are the solutions to the following ODEs both subjected to zero flux boundary conditions

$$\frac{\epsilon[c^o(x,t+\Delta t)-c(x,t)]}{\Delta t}=\nabla[\epsilon D^{\text{eff}}\nabla c^o(x,t+\Delta t)], \tag{3.17}$$

$$\frac{\epsilon c^j(x,x_0,t+\Delta t)}{\Delta t}=\nabla[\epsilon D^{\text{eff}}\nabla c^j(x,x_0,t+\Delta t)]+\delta(x-x_0), \tag{3.18}$$

with $\delta(x-x_0)$ being the Dirac delta function.

At this point, solutions to the solid and electrolyte phase diffusion equations are decoupled from the over-potential and pore wall flux since in solving Eqs. (3.12), (3.14), (3.17) and (3.18), η and j_{vol} are not needed. Moreover, for the cases when D_s, D^{eff} and ϵ are constant, analytic solutions are available for Eqs. (3.14) and (3.18) [194]. c_s^j and c^j can also be used to obtain a linearized form for the terms including $V(\eta)$ and $\ln c(\eta)$ on the RHS of Eq. (3.6), respectively. In conclusion, one needs to solve the decoupled Eqs. (3.12), (3.14), (3.17), (3.18), and (3.6) instead of solving the coupled system of non-linear equations in Table 3.1 for LIB modeling. In addition, if a constant time step is used, Eqs. (3.14) and (3.18) need to be solved once at the beginning of the simulation. The only major approximation in deriving this decoupled formulation is the finite time differencing which is inevitable in numerical simulations. An extended version of this decoupled formulation, which includes SEI resistance and double layer capacitive effects coupled with a thermal model, is given in [196, 195]. Interested reader is referred to [194] for more details on linearizing the over-potential equation (3.6), grid generation and the time evolution strategy.

3.3 Non-intrusive Polynomial Chaos Expansion

In the following, we will review the least squares regression and the compressive sampling methods which are used to compute the PC coefficients in a non-intrusive fashion, i.e., via sampling.

3.3.1 Least squares regression

The least squares regression technique is basically the regression of the exact solution $u(\boldsymbol{\xi})$ in the PC bases [237]. Given the set of samples $\{\boldsymbol{\Xi}^{(i)}\}_{i=1}^{N}$, generated randomly, for instance, according to $\rho(\boldsymbol{\xi})$, and

the corresponding solution realizations $\{u(\boldsymbol{\Xi}^{(i)})\}_{i=1}^{N}$, the discrete representation of (2.6) can be written as

$$\boldsymbol{\Psi}\boldsymbol{\alpha} \approx \boldsymbol{u}, \tag{3.19}$$

where $\boldsymbol{u} = (u(\boldsymbol{\Xi}^{(1)}), \cdots, u(\boldsymbol{\Xi}^{(N)}))^T \in \mathbb{R}^N$ contains the realizations of the QoI, $\boldsymbol{\Psi}(i,j) = \psi_j(\boldsymbol{\Xi}^{(i)}) \in \mathbb{R}^{N\times P}$ is the measurement matrix containing samples of the PC basis, and $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_P)^T \in \mathbb{R}^P$ is the vector of PC coefficients.

PC coefficients $\boldsymbol{\alpha}$ may be approximated by solving the least squares problem

$$\min_{\boldsymbol{\alpha}} \|\boldsymbol{u} - \boldsymbol{\Psi}\boldsymbol{\alpha}\|_2, \tag{3.20}$$

where $\|\cdot\|_2$ is the l_2 norm. When $\boldsymbol{\Psi}$ is full rank, the solution to (4.4) is computed from the normal equation

$$(\boldsymbol{\Psi}^T\boldsymbol{\Psi})\boldsymbol{\alpha} = \boldsymbol{\Psi}^T\boldsymbol{u}. \tag{3.21}$$

In general, a stable solution $\boldsymbol{\alpha}$ to (4.5) requires $N > P$ realizations of u. In [107, Theorems 2.2 and 3.1], it is shown that for d-dimensional Legendre polynomials of total order p, for the case of $d > p$, a stable solution recovery from (4.4) can be guaranteed with a number of samples given by

$$N \geq 3^p \, C \, P \log(P), \tag{3.22}$$

where C is an absolute constant. This suggests that the number of samples N depends linearly on P (up to a logarithmic factor) for the least squares regression method.

For high dimensional complex problems, such as LIBs, generating $N > P$ realizations may be computationally expensive. In such cases, when the solution $u(\boldsymbol{\xi})$ depends *smoothly* on $\boldsymbol{\xi}$, the PC coefficients are often sparse, i.e., many of them are negligible. In these cases, CS may be employed to compute the coefficients with $N < P$ realizations [59, 107, 184, 108, 211, 262, 124, 264]. Specifically, the importance sampling approach of [108] ensures an accurate computation of $\boldsymbol{\alpha}$ with a number of solution realizations that depends linearly (up to a logarithmic factor in P) on the number of dominant coefficients. We next review the CS approach, which we use later for the UQ of LIBs.

3.3.2 Compressive sampling

Compressive sampling/sensing is an emerging direction in signal processing which enables (up to) exact reconstruction of signals admitting sparse representations with a small number of signal measurements [56, 29, 36]. It was first introduced to the UQ field in 2011 by Doostan and Owhadi [59] where they used CS to approximate sparse PC solutions to stochastic PDEs. The main requirement in CS methods is to have a sparse solution at the stochastic level, that is a small fraction of PC coefficients in (2.6) are dominant and contribute to the solution statistics. The ultimate goal of CS is to approximate the sparse PC coefficients $\boldsymbol{\alpha}$ accurately and robustly with $N < P$ realizations of $u(\boldsymbol{\xi})$.

With $N < P$, the underdetermined linear system in (3.19) is ill-posed and generally has infinitely many solutions. Sparsity of the PC coefficients $\boldsymbol{\alpha}$ allows a regularization of (3.19) to ensure a well-posed solution [59]. This can be achieved by solving the Basis Pursuit Denoising (BPDN) problem

$$\min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_1 \quad \text{subject to} \quad \|\boldsymbol{\Psi}\boldsymbol{\alpha} - \boldsymbol{u}\|_2 \leqslant \gamma. \tag{3.23}$$

Minimization of the l_1 norm in (3.23) promotes sparsity in $\boldsymbol{\alpha}$, while the l_2 residual norm controls the accuracy of the truncated PC expansion with the tolerance γ. Several numerical techniques are available in the literature to solve the BPDN problem [59]. Among those, we adopt the Spectral Projected Gradient algorithm (SPGL1) of [17] implemented in the `SPGL1 package` for MATLAB [16].

The accuracy of $\boldsymbol{\alpha}$ computed from (3.23) depends on the sample size N and the truncation error $\|\boldsymbol{\Psi}\boldsymbol{\alpha} - \boldsymbol{u}\|_2$ in (3.23). In general, the truncation error may be decreased by increasing the order p of the PC basis, which leads to a larger number of coefficients, P. This, in turn, requires a larger number of samples N to maintain the stability of the BPDN problem. The minimum sampling rate depends on the type of the PC basis, the sparsity of $\boldsymbol{\alpha}$, and the sampling distribution according to which $\{\boldsymbol{\Xi}^{(i)}\}_{i=1}^{N}$ are generated, and is shown to linearly depend on the number of dominant PC coefficients [108, Theorems 3.1 and 4.2]. More precisely, for d-dimensional Legendre polynomials of total order p, where $d > p$, the number of samples N required to guarantee a stable solution recovery from (3.23) is given by

$$N \geq 3^p \, C \, S \log(P), \tag{3.24}$$

where S is the number of dominant coefficients. In (3.24), N depends primarily on S and weakly on P through the $\log(P)$ term. For high dimensional complex problems, we usually have $P \gg S$, hence, a comparison of (3.22) and (3.24) suggests that the CS approach requires considerably smaller number of samples in comparison to the least squares regression.

In practice, one may start by approximating a lower order PC expansion when N is small and increase p when a larger number of samples become available. Another important factor in the sparse reconstruction is the selection of the truncation error tolerance γ. The ideal value for the tolerance is $\gamma \approx \|\boldsymbol{\Psi}\boldsymbol{\alpha}_{\text{exact}} - \boldsymbol{u}\|_2$. Since the exact PC coefficients $\boldsymbol{\alpha}_{\text{exact}}$ are not known, selecting larger values than $\|\boldsymbol{\Psi}\boldsymbol{\alpha}_{\text{exact}} - \boldsymbol{u}\|_2$ for γ deteriorates the accuracy of the approximation, while smaller choices may result in over-fitting the solution samples and, thus, less accurate results. In the numerical results of Section 3.6, we employ the cross-validation approach in [59, Section 3.5] to optimally choose γ. For the sake of completeness, this cross-validation approach is summarized in Algorithm 3. For more details about the CS method, the interested reader is referred to [59, 108].

Algorithm 3 Algorithm for cross-validation estimation of γ.

1: Divide the N solution samples to N_r reconstruction and N_v validation samples.
2: Choose multiple values for γ_r such that the exact truncation error $\|\boldsymbol{\Psi}\boldsymbol{\alpha}_{\text{exact}} - \boldsymbol{u}\|_2$ of the reconstruction samples is within the range of γ_r values.
3: For each value of γ_r solve the BPDN problem (3.23) using the N_r reconstruction samples to compute $\boldsymbol{\alpha}_r$.
4: For each value of γ_r, compute the truncation error $\gamma_v := \|\boldsymbol{\Psi}_v\boldsymbol{\alpha}_r - \boldsymbol{u}_v\|_2$ of the N_v validation samples.
5: Find the minimum value of γ_v and and its corresponding $\hat{\gamma}_r := \gamma_r$.
6: Set $\gamma = \sqrt{\frac{N}{N_r}}\hat{\gamma}_r$.

3.4 Global Sensitivity Analysis

Identification of the most important random inputs affecting the variations in the cell voltage, capacity, and concentrations is one of the objectives of this study. This is achieved by performing a global sensitivity analysis (SA) to quantify the specific effects of random inputs on the variance of the QoI.

Among the available techniques to perform global SA, we use the Sobol' indices [228] which are widely used due to their generality and accuracy. Sudret [237] introduced an analytic approach to compute

the Sobol' indices as a post-processing of the PC coefficients. Let us assume the PC coefficients in (2.6) are computed. The first order PC-based Sobol' index S_k, which represents the sole effects of the random input ξ_k on the variability of $u(\boldsymbol{\xi})$, is given by

$$S_k = \sum_{\boldsymbol{i} \in \mathscr{I}_k} \alpha_{\boldsymbol{i}}^2 / \mathrm{var}[u], \quad \mathscr{I}_k = \{\boldsymbol{i} \in \mathbb{N}_0^d : i_k > 0, i_{m \neq k} = 0\}, \tag{3.25}$$

where $\mathrm{var}[u]$ is given in (1.7). In computing S_k, it is assumed that all random inputs except ξ_k are fixed, therefore, S_k does not represent the effects of the interactions between ξ_k and other random inputs. In order to quantify the total effects of the random input ξ_k, including the interactions between random inputs on the variability of $u(\boldsymbol{\xi})$, one needs to compute the total PC-based Sobol' indices defined as

$$S_k^T = \sum_{\boldsymbol{i} \in \mathscr{I}_k^T} \alpha_{\boldsymbol{i}}^2 / \mathrm{var}[u], \quad \mathscr{I}_k^T = \{\boldsymbol{i} \in \mathbb{N}_0^d : i_k > 0\}. \tag{3.26}$$

The smaller S_k^T, the less important random input ξ_k. For the cases when $S_k^T \ll 1$, ξ_k is considered as insignificant and may be replaced by its mean value without considerable effects on the variability of $u(\boldsymbol{\xi})$. In this study, we employ S_k^T as a measure to identify the most important random inputs of the LIB model considered.

3.5 Uncertainty in LIB Model Parameters

The model parameters of LIBs are measured experimentally and are accompanied by uncertainty due to natural or experimental variability. Some of these parameters are measured using complex electrochemical techniques, while others are obtained via simple experiments. In the following, we will discuss a number of such techniques. It should be noted that because of the limited data available in the literature, we could not avoid making assumptions on the probability distribution for some of the parameters. Additionally, the reported uncertainty models are specific to the LiC_6/$LiCoO_2$ cell we consider here.

3.5.1 Porosity, ϵ

Porosity is defined as the ratio of the pore volume to the bulk volume. Manufacturers choose the porosity as a trade off between power and energy, i.e., the higher the porosity, the higher the power and

the lower the capacity [216]. There are several methods to measure the porosity, such as the Method of Standard Porosimetry (MSP) [66], porosity measurement using liquid or gas absorption methods according to the American Society for Testing and Materials (ASTM) D-2873 [270], and X-ray tomography [44].

In [208], a uniform distribution [0.28, 032] is considered for the porosity of $LiCoO_2$ based on experimental data, which has $\pm 6.7\%$ (of the mean) variation around the mean. DuBeshter et al. [66] reported $\pm 4.5\%$ variability for porosity of the graphite anode electrode, while a higher value of $\pm 12.6\%$ is reported for the separator. Hence, in the present study, we assume a uniform distribution for the porosity with $\pm 6.7\%$, $\pm 4.5\%$, and $\pm 12.6\%$ variation around the mean in $LiCoO_2$ cathode, LiC_6 anode, and separator, respectively. Defining the coefficient of variation (COV) as the ratio of the standard deviation to the mean, the assumed variations translate to COVs of 0.026, 0.073, and 0.038 in anode, separator, and cathode, respectively.

3.5.2 Solid particle size, r_s

The flux of Li^+ at the electrode-electrolyte interface is affected by the solid particle size in electrodes, as it defines the available surface area for the reaction. The maximum battery power may be increased by decreasing the particle size of the electrode material and increasing the surface area per volume.

The particle size distribution may be measured by a laser diffraction and scattering device [43], X-ray computed tomography (XCT) [128], or focused ion beam tomography [215].

Santhanagopalan and White [207] considered a normal distribution for a graphite anode with the mean and standard deviation of 6.2 μm and 0.42 μm, respectively, to quantify the effects of random particle size. For the particle size distribution of $LiCoO_2$, a normal distribution with mean of 7.7 μm and standard deviation of approximately 1.5 μm is reported in [150]. The corresponding COV values are 0.0677 and 0.1948 for the graphite anode and $LiCoO_2$ cathode, respectively. Nominal values, i.e., mean, of the particle sizes in both electrodes are equal to 2.0 μm in our cell. Because of the limited data available in the literature on the particle size distribution of our specific electrodes, we use these COVs to find the corresponding standard deviations. Based on these COVs, we assume a normal distribution with a mean of 2.0 μm and standard deviation of 0.1354 μm for the graphite anode. For the cathode electrode, the same mean value but a standard deviation of 0.3896 μm is considered.

Remark 3. *When one draws samples from a Gaussian distribution for the solid particle size r_s, one should assure that all the r_s samples are strictly positive. For this purpose, we here employed a truncated Gaussian distribution bounded between mean ± 3 standard deviation of r_s.*

3.5.3 Bruggeman coefficient, brugg

In porous electrode theory, instead of specifying the exact position and shapes of all pores and particles, a volume-averaged formulation is used [240]. In this model, the effective transport properties of the liquid phase, D^{eff} and κ^{eff}, are obtained using the porosity ϵ and tortuosity τ via

$$\kappa^{\text{eff}} = \frac{\epsilon\kappa}{\tau}, \qquad D^{\text{eff}} = \frac{\epsilon D}{\tau}. \tag{3.27}$$

Tortuosity τ is a geometric parameter and depends on the porous electrode structure [242]. Although in recent years researchers have attempted to measure τ via experimental techniques [242, 69, 66, 216], because of the experimental complexities, τ has been commonly used as a model parameter that is calibrated by experiments [63]. More precisely, τ is computed via the well-known Bruggeman relation

$$\tau = \epsilon^{(1-\mathrm{brugg})}, \tag{3.28}$$

where the Bruggeman exponent brugg is chosen to match numerical results with experimental data, and is usually assumed to be 1.5 [242].

Since the Bruggeman relation is widely used in LIB simulations, instead of taking τ as a random input parameter, we choose the Bruggeman exponent brugg to be an uncertain parameter in determining the effective properties. In [66], a uniform distribution with $\pm 4.9\%$ of the mean variation around the mean is reported for the Bruggeman exponent of the LiC_6 anode, while $\pm 33.3\%$ is considered for the Bruggeman exponent in separator brugg_s [208]. For this study, we assume a uniform distribution for the Bruggeman exponent with $\pm 5.0\%$ variations around the mean in both electrodes and $\pm 20.0\%$ in separator, which corresponds to COVs of 0.029 and 0.115, respectively.

3.5.4 Li^+ transference number, t_+^0

When an LIB is discharged, electrolyte salt dissociates into Li^+ and PF_6^- ions. The portion of the current that is carried by Li^+ ions is called the Li^+ transference number. The higher the Li^+ transport number, the higher the battery power [273].

Hittorf's method [15], ac impedance spectroscopy [129], or pulsed field gradient NMR (pfg-NMR) technique [85] are used to measure the transference number in electrolytes. The experimental error of pfg-NMR technique for binary solutions is estimated to be around 5.0% [247]. Hence, we let the transference number uniformly change between 0.345 and 0.381 in this study.

3.5.5 Salt diffusion coefficient in the liquid phase, D

The salt diffusion coefficient D is a measure of the friction forces between the ions and the solvents [180]. In order to restrict the performance-limiting salt concentration gradients, which form in the electrolyte during polarization, it is critical to have a high salt diffusion coefficient [169]. High values of D lead to higher battery power. Experimental methods such as cyclic voltammetry (CV) [263] and electrochemical impedance spectroscopy (EIS) [251] have been used to measure D.

Salt diffusion coefficient is usually reported as a constant, but in [247] it is assumed to be a function of temperature and salt concentration. In this study, we assume a uniform distribution for D with $\pm 10.0\%$ variation around the mean, which corresponds to COV of 0.0577.

3.5.6 Diffusion coefficient of the solid, D_s

Diffusion coefficient of the solid phase plays an important role in the performance of LIBs since it affects the intercalation flux. Electrochemical techniques such as CV, GITT and EIS have been used to determine D_s [257]. D_s has been mostly treated as a constant value in LIB simulations [90], while it has been shown that D_s depends on the intercalation level, i.e., the ratio of $c_s^{\text{surf}}/c_{s,max}$ [141]. The values of D_s for the same materials, reported by different research groups, may differ by several orders of magnitude. These large differences may suggest that a Fickian diffusion model which is used to calibrate D_s does not

correctly describe the transport of Lithium within the active particles and other transport models need to be considered.

In this study, we assume that D_s does not depend on the intercalation level. Hence, for D_s, we assume a uniform distribution with the same COV we assumed for D.

3.5.7 Electronic conductivity of the solid, σ

Capacity of LIBs may be improved by increasing the solid phase electronic conductivity by using conductive additives in the electrode materials [40]. Two-point and four-point probe techniques have been used to measure the electronic conductivity of electrode materials [241]. Although σ depends on temperature and state of the charge [210], it is mostly treated as a constant in LIB simulations.

Reported values of σ in the literature for LiC_6 anode are mostly around 100 $\mathrm{S \cdot m^{-1}}$ [236, 61, 90], while the electronic conductivity of $LiCoO_2$ cathode is reported to be 100 $\mathrm{S \cdot m^{-1}}$ [236] and 10 $\mathrm{S \cdot m^{-1}}$ [61]. In our LIB model, the nominal values of σ is equal to 100 $\mathrm{S \cdot m^{-1}}$ for both electrodes. Because of the limited available data in the literature on the measurement uncertainties in σ, we assume that in both electrodes it changes uniformly with $\pm 10.0\%$ variation.

3.5.8 Reaction rate constant, k

Exchange current density i_{ex} is measured at the initial state of the battery [185]. Using the relation $i_{ex} = Fk(c_s^{\mathrm{surf}})^{0.5}(c_{s,max} - c_s^{\mathrm{surf}})^{0.5}(c)^{0.5}$, one can calculate the reaction rate constant k for each electrode at the initial values of c_s^{surf} and c. Although k shows an Arrhenius type dependence on temperature and depends on the nature of the electrode surface [170], it is usually treated as a constant in LIB simulations. Higher reaction rate constants are favored for Li-ion batteries since the reaction is more reversible, and polarization effects are lower.

Because of the lack of experimental data on the reaction rate constant, we assumed that k has a uniform distribution with $\pm 10.0\%$ variation around the mean in both electrodes.

A list of random parameters considered in this study is presented in Table 3.2. In addition to the discussed input parameters, we also took the lengths L of electrodes and separator to be random in order to

study the effect of uncertainties in geometrical parameters of the battery. It also should be noted that in our battery model, the effective ionic conductivity of the liquid phase κ^{eff} is considered to be a function of liquid concentration. Hence, κ^{eff} will be automatically a random parameter since the porosity is considered to be random. We note that the anodic and cathodic transfer coefficients (considered to be 0.5 in this study) as well as the open circuit potentials are also subjected to uncertainties; however, we treat them as deterministic parameters in this study.

Table 3.2: List of random and deterministic LIB inputs used in this study.

Random Input	**Nominal Value**	**Distribution**
$r_{s,a}$ [μm]	2	Gaussian, $\mu = 2$, $\sigma = 0.1354$
$r_{s,c}$ [μm]	2	Gaussian, $\mu = 2$, $\sigma = 0.3896$
ϵ_a	0.485	Uniform, [0.46, 0.51]
ϵ_s	0.724	Uniform, [0.63, 0.81]
ϵ_c	0.385	Uniform, [0.36, 0.41]
brugg_a	4	Uniform, [3.8, 4.2]
brugg_s	4	Uniform, [3.2, 4.8]
brugg_c	4	Uniform, [3.8, 4.2]
t_+^0	0.363	Uniform, [0.345, 0.381]
D [$\text{m}^2 \cdot \text{s}^{-1}$]	7.5×10^{-10}	Uniform, [6.75, 8.25] $\times 10^{-10}$
$D_{s,a}$ [$\text{m}^2 \cdot \text{s}^{-1}$]	3.9×10^{-14}	Uniform, [3.51, 4.29] $\times 10^{-14}$
$D_{s,c}$ [$\text{m}^2 \cdot \text{s}^{-1}$]	1×10^{-14}	Uniform, [0.9, 1.1] $\times 10^{-14}$
σ_a [$\text{S} \cdot \text{m}^{-1}$]	100	Uniform, [90, 110]
σ_c [$\text{S} \cdot \text{m}^{-1}$]	100	Uniform, [90, 110]
k_a [$\text{m}^4 \cdot \text{mol} \cdot \text{s}$]	5.03×10^{-11}	Uniform, [4.52, 5.53] $\times 10^{-11}$
k_c [$\text{m}^4 \cdot \text{mol} \cdot \text{s}$]	2.334×10^{-11}	Uniform, [2.10, 2.56] $\times 10^{-11}$
L_a [μm]	80	Uniform, [77, 83]
L_s [μm]	25	Uniform, [22, 28]
L_c [μm]	88	Uniform, [85, 91]

3.6 Numerical Example

In this section, we present a numerical example to demonstrate the application of our proposed UQ approach to LIBs. The one-dimensional $LiC_6/LiCoO_2$ cell we consider here is studied in [236]. We present our results for three different discharge rates of 0.25C, 1C, and 4C to study its effects on the propagation of uncertainty. We note that assuming a constant discharge rate is an idealized scenario. Typically, the battery loading varies over the course of the discharge process, depending on an application-dependent usage of the battery. Hence, in addition to the previously mentioned uncertain parameters, the discharge rate should also be considered as an uncertain input for the LIB model. For the sake of comparison, we treat the discharge rate as a deterministic input in this study, while our proposed UQ framework can incorporate

random discharge rates with no difficulties, perhaps, with a cost of running additional battery simulations.

For the spatial discretization of the LIB governing equations, we used the Finite Difference Method on the non-uniform grids described in [194]. We also performed a mesh convergence analysis to ensure that spatial discretization errors are inconsequential.

Remark 4. *Working with a fine mesh and a small time step is advised when one employs sampling-based UQ techniques since a coarse discretization may not return a converged/accurate solution for all samples. In our calculations, we found that having 120 non-uniform grid points in each region and 26 non-uniform grid points in the spherical solid particles are sufficient to obtain accurate realizations for all of the samples. Moreover, constant time steps of 1.0, 0.1, and 0.05 seconds were used for 0.25C, 1C, and 4C rates of discharge, respectively.*

The cell model incorporating the nominal values in Table 3.2 is referred to as the *deterministic* model, while the *stochastic* model uses the distributed input parameters of Table 3.2. The number of random inputs of our stochastic LIB model is $d = 19$, which may be considered *high*. We use the CS technique in this study to approximate the sparse PC coefficients since, in comparison to least squares regression, CS requires smaller number of battery simulations.

Having at hand a deterministic LIB solver, the next step in sampling-based PC expansion is to generate N realizations of model parameters listed in Table 3. To this end, we assign an independent random variable ξ_k, $k = 1, \ldots, d$, to each parameter, and consider an appropriate linear transformation of each ξ_k to match the PDF of the corresponding parameter in Table 3.2. For Gaussian and uniform PDFs, we use ξ_k's that are standard Gaussians and uniformly distributed between $[-1, 1]$, respectively. The independent samples of $\boldsymbol{\xi}$, i.e., $\{\boldsymbol{\Xi}^{(i)}\}_{i=1}^{N}$, are used to generate N independent samples of the model parameters, for which the LIB model is simulated to obtain N realizations of the output QoIs, $\{u(\boldsymbol{\Xi}^{(i)})\}_{i=1}^{N}$. In our numerical experiments, we simulate each battery realization until a cut-off potential of 2.8 V is reached. Then using $\{\boldsymbol{\Xi}^{(i)}\}_{i=1}^{N}$ and $\{u(\boldsymbol{\Xi}^{(i)})\}_{i=1}^{N}$, we solve the BPDN problem in (3.23) to approximate the vector of PC coefficients $\boldsymbol{\alpha}$. These in turn will be used to compute the statistics of $u(\boldsymbol{\xi})$, such as the mean and variance given in (1.6) and (1.7), respectively, as well as the total Sobol' indices defined in (3.26).

We find that a PC expansion of order $p = 3$ and $N = 1000$ battery simulations are needed to achieve a *validation error* smaller than 1.0% as specified next. Following the cross-validation procedure described in [59, Section 3.5], we divide the $N = 1000$ samples into $N_r = 900$ reconstruction and $N_v = 100$ validation samples to estimate the optimum value of γ in (3.23), using which we compute the solution $\boldsymbol{\alpha}$ from (3.23). To verify the accuracy of the resulting PC expansion, we compute the validation error

$$\text{relative error} = \frac{\|\boldsymbol{u}_v - \boldsymbol{\Psi}_v \boldsymbol{\alpha}_r\|_2}{\|\boldsymbol{u}_v\|_2}, \tag{3.29}$$

where $\boldsymbol{u}_v$ is the vector of $N_v = 1000$ additional realizations of QoI (not used in computing $\boldsymbol{\alpha}$) and $\boldsymbol{\Psi}_v$ is the measurement matrix corresponding to $\boldsymbol{u}_v$. Stated differently, the error in (4.23) determines the accuracy of constructed PC model in predicting independent QoI realizations. In practice, there is no need to generate additional samples for validation and we only used a larger number of validation samples to demonstrate the accuracy and robustness of our proposed method.

Remark 5. *In practice, we rely on the validation error to estimate the solution accuracy and decide to possibly increase the total order of PC expansion p. When the validation error does not reduce by increasing the number of numerically generated LIB samples for a given expansion order p, one may achieve smaller validations errors by increasing p with a cost of increased number of LIB simulations.*

Algorithm 4 summarizes the main steps in our proposed UQ framework for LIBs for a fixed PC expansion order p.

Algorithm 4 Summary of the main steps in the proposed UQ framework for LIBs.

1: Identify uncertain parameters and their probability distributions (Section 3.5).
2: **while** the validation error in (4.23) is larger than a user-defined threshold **do**
3: Generate N realizations of the uncertain parameters based on their PDFs.
4: Perform deterministic LIB simulations to obtain the output QoI (Section 3.2).
5: Evaluate the corresponding realization of the PC basis (Section 3.3).
6: Perform the cross-validation approach in Algorithm 3 to optimally estimate γ.
7: Solve the BPDN problem in (3.23) to approximate the PC coefficients.
8: Increase N.
9: **end while**
10: Compute the statistics of the output QoI via (1.6) and (1.7) and the total Sobol' indices via (3.26).

3.6.1 Effects of input uncertainties on cell capacity

Capacity is defined as the available energy stored in a fully charged LIB and is one of the most important factors affecting the battery performance. However, effects of LIB model uncertainties on capacity estimation are not yet fully explored.

Fig. 3.2 demonstrates the effects of LIB model uncertainties in estimating the cell capacity. In Fig. 3.2(a), cell capacities as a function of the cell voltage obtained from deterministic and stochastic models are presented for 0.25C, 1C and 4C rates of galvanostatic discharge. Shaded areas around the mean of stochastic cell capacity represent three standard deviation intervals. As it can be seen, noticeable deviation from the deterministic capacity starts when the cell voltage experiences a rapid drop-off. As the discharge rate is decreased, the onset of this deviation happens at higher cell voltages, i.e., at $\phi_{\text{cell}} \approx 3.7$ V for $I = 0.25$C and $\phi_{\text{cell}} \approx 3.3$ V for $I = 1$C. This plot shows that for low to medium rates of discharge, a deterministic simulation of LIB overestimates the cell capacity. For 4C discharge rate, the mean of stochastic cell capacity overlaps with the cell capacity predicted by the deterministic model. One may suspect that at high discharge rates, input uncertainties have no significant effect on the cell capacity, but the probability bounds around the mean of cell voltage for $I = 4$C suggest otherwise. As it can be seen, the largest standard deviation of the cell capacity at the end of discharge is for the 4C discharge rate.

Fig. 3.2(b) shows the validation error in approximating the PC coefficients of the cell capacity. With $p = 3$ and $d = 19$, the number of PC coefficients $P = |\mathscr{I}_{19,3}| = 1540$, which is larger than $N = 1000$. The accurate solution approximation in this under-sampled setting is achieved by minimizing the l_1 norm of $\boldsymbol{\alpha}$ in the BPDN problem (3.23).

A comparison of deterministic (solid line) and distributed (dashed line) cell capacities is also presented in Fig. 3.2(c). This plot suggests that the probability of achieving deterministic cell capacity when the cell is subjected to assumed input uncertainties is very small for low to medium discharge rates, while this probability is considerably larger for higher discharge rates.

An alternative sampling-based approach to compute the PC coefficients is the stochastic collocation (SC) method [153, 259, 11]. The main idea behind the SC technique is to sample the output QoI at particular

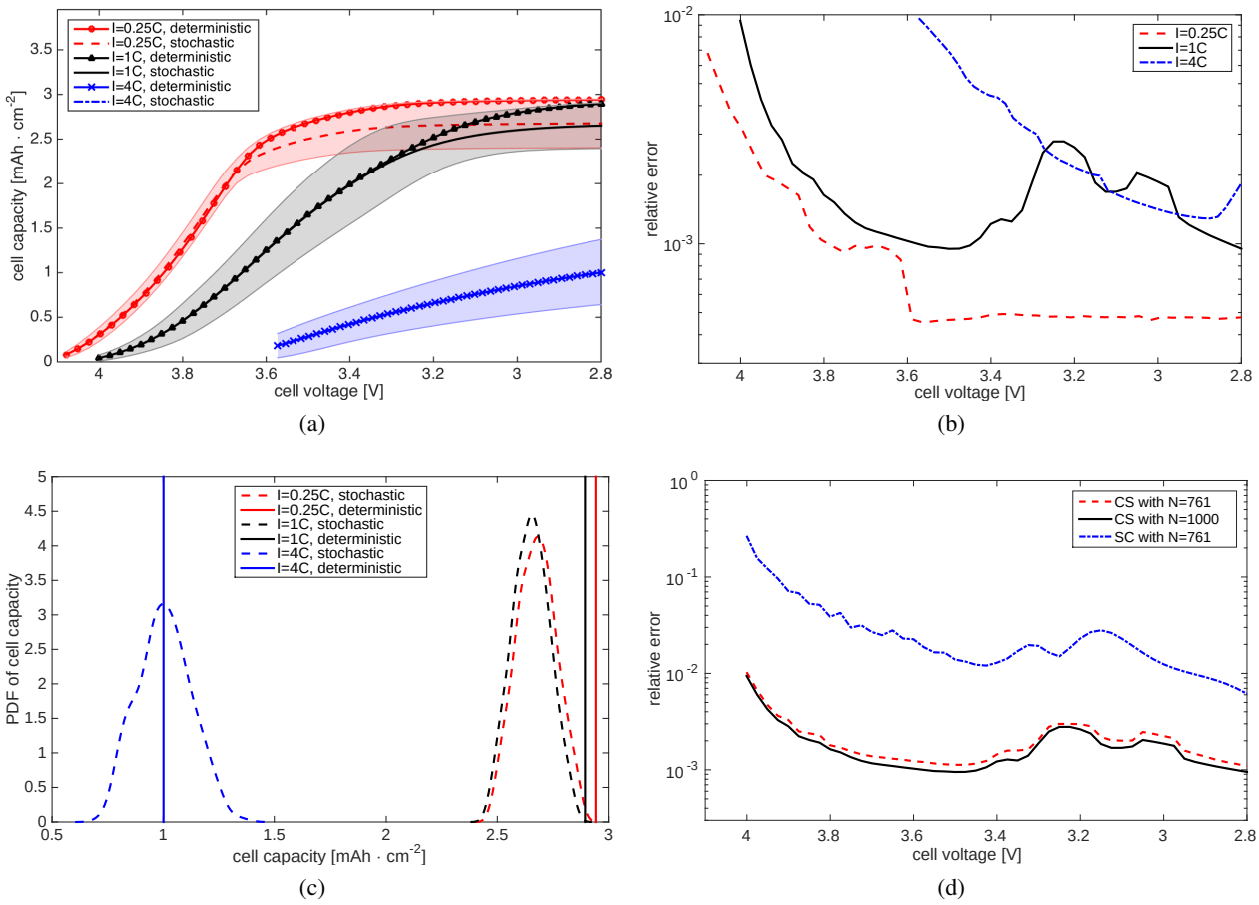


Figure 3.2: Comparison of stochastic and deterministic battery models for $I = 0.25$C, 1C and 4C rates of galvanostatic discharge. (a) Stochastic and deterministic cell capacities as a function of the cell voltage. The shaded areas are probability bounds of three standard deviations around the mean; (b) Relative error as a function of the cell voltage; (c) PDFs of the cell capacity; (d) Comparison of the relative error obtained by the CS method and the level two SC technique on a Clenshaw-Curtis grid [175] for 1C rate of discharge. Level three SC method requires 9976 LIB simulations.

points in the stochastic space and then approximate the solution via interpolation or the solution statistics by numerical integration. For high-dimensional problems, sparse tensor products first introduced by Smolyak [226], which may be built upon the Clenshaw-Curtis abscissas, has been proposed to relax the *curse-of-dimensionality* associated with full tensor products [175]. The accuracy of this method is controlled by the so-called level parameter. The number of required LIB simulations increases with the magnitude of the level parameter. SC may also be used to compute the PC coefficients via the projection equation $\alpha_{\boldsymbol{i}} = \mathbb{E}[u(\cdot)\psi_{\boldsymbol{i}}(\cdot)]/\mathbb{E}[\psi_{\boldsymbol{i}}^2]$. In particular, the numerator in this equation is a d-dimensional integral over Ω which

may be computed using tensor product quadrature or cubature rules. Detailed description of these techniques is beyond the scope of this paper and the interested reader is referred to the provided references.

In this study, we employ the SC technique with sparse tensor products based on Clenshaw-Curtis abscissas to compare with the proposed CS approach. For the level parameter of two, one needs 761 LIB simulations in the SC method. Fig. 3.2(d) compares the validation error in approximating the PC coefficients of the cell capacity obtained by SC and CS methods for the cell discharged at 1C rate. As it can be seen, CS with 761 samples results in the validation errors that are smaller (about one order of magnitude) than those reported by the SC technique. Increasing the level parameter of the SC method to three, in order to increase the accuracy, requires 9976 LIB simulations. A comparison between the level two SC and CS approaches confirms the advantage of our proposed UQ framework, hence, we did not perform level three SC due to the high computational costs. Since we use random samples in CS, unlike in the SC technique, the number of samples are not dictated by the method and one may freely choose a minimum number of additional LIB simulations to achieve the desired accuracy.

As discussed in Section 3.4, we compute the total Sobol' indices given in (3.26) to identify the most important random inputs. We assume that a random input ξ_k with a maximum total Sobol' index S_k^T smaller than 0.01 has no significant effects on the variability of the output QoI and may be treated as a deterministic input. This selection criteria is problem dependent and may be changed depending on the application requirements. Fig. 3.3 shows the total Sobol' indices associated with the cell capacity over a course of discharge for $I = 0.25$C, 1C and 4C. In this paper, we only present significant Sobol' indices, i.e., those with $\max(S_k^T) > 0.01$, in the global SA plots. Moreover, the legends of the Sobol' index plots are sorted such that the first legend corresponds to the largest S_k^T, while the last legend represents the smallest Sobol' index. The following observations from Fig. 3.3 are worthwhile highlighting:

- Independent of the discharge rate, porosity ϵ and Bruggeman coefficient brugg in anode, separator, and cathode are among the most important random inputs. In other words, variation in the the cell capacity is highly affected by uncertainty in tortuosity τ.

- For all three discharge rates, Fig. 3.3 shows that σ_a, σ_c, t_+^0, $D_{s,a}$, $D_{s,c}$, k_a, k_c, and $r_{s,a}$ are

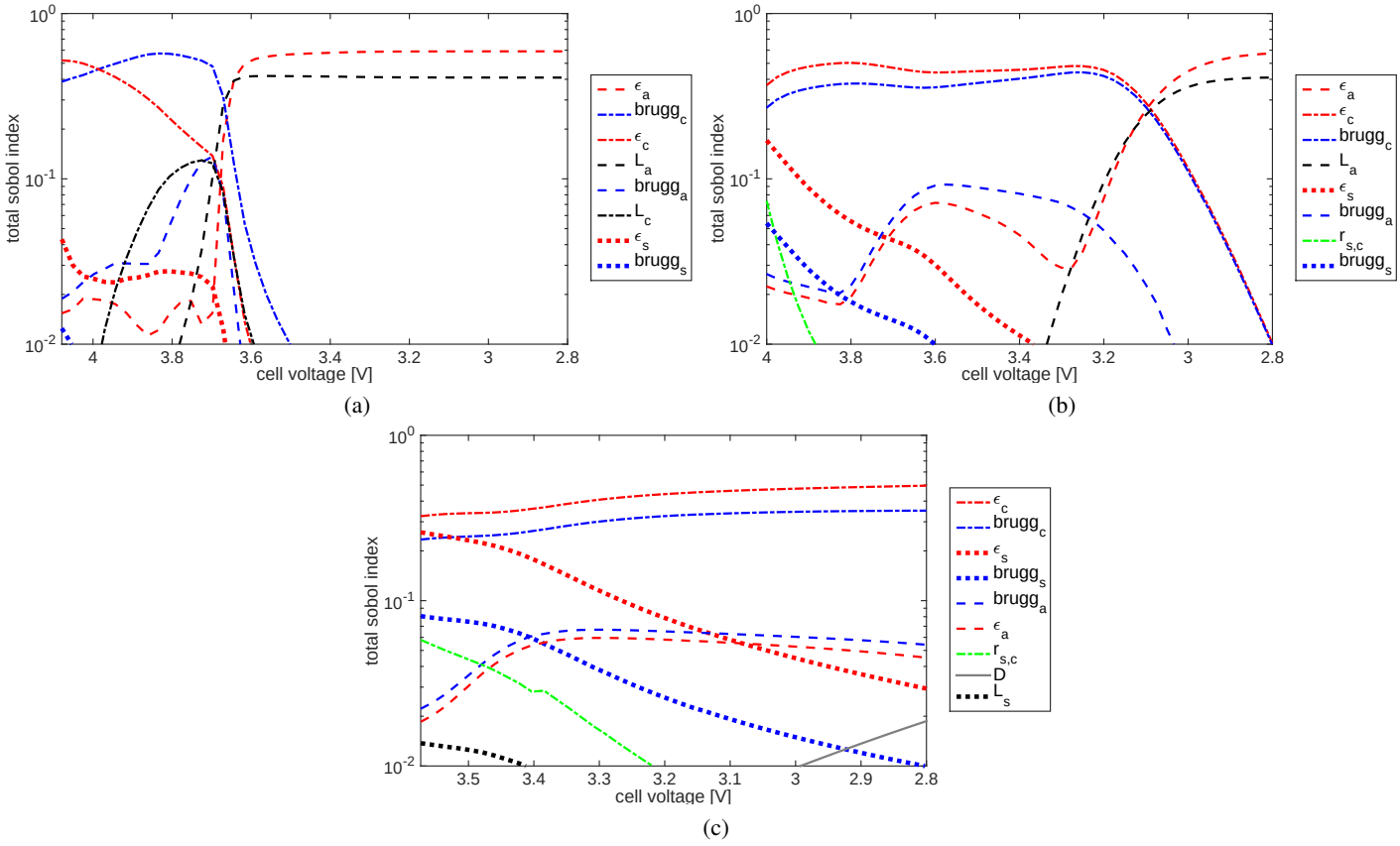


Figure 3.3: Global sensitivity analysis of the cell capacity for various discharge rates: (a) $I = 0.25$C; (b) $I = 1$C; (c) $I = 4$C.

insignificant random inputs and their uncertainties have no important effects on the variability of the cell capacity. Consequently, expensive and accurate quality control measures for these parameters are not required when one aims at reducing the variations in cell capacity.

- As the discharge rate increases, uncertainties in the length of electrodes, i.e., L_a and L_c, become less important, while the effects of variability in the length of separator L_s are more pronounced for $I = 4$C.

- For $I = 4$C, the diffusion coefficient of the liquid phase D is an important random input, while for low to medium rates, its corresponding Sobol' index S_D^T is smaller than 0.01.

- By increasing the discharge rate, the number of important random inputs at the end of discharge

increases; two for $I = 0.25$C, four for $I = 1$C, and seven for $I = 4$C. In other words, for high discharge rate LIB applications, accurate quality control measures are needed for a larger number of LIB parameters.

- For low discharge rates, variability in the solid particle size r_s has no significant effects on the variations of the cell capacity.

In general, Fig. 3.3 suggests that in determining the most significant random inputs on the variations of the LIB capacity, considering the discharge rate is crucial. Again, we emphasize that these observations may change when more accurate LIB models that account for cell degradations and other physical phenomena are employed.

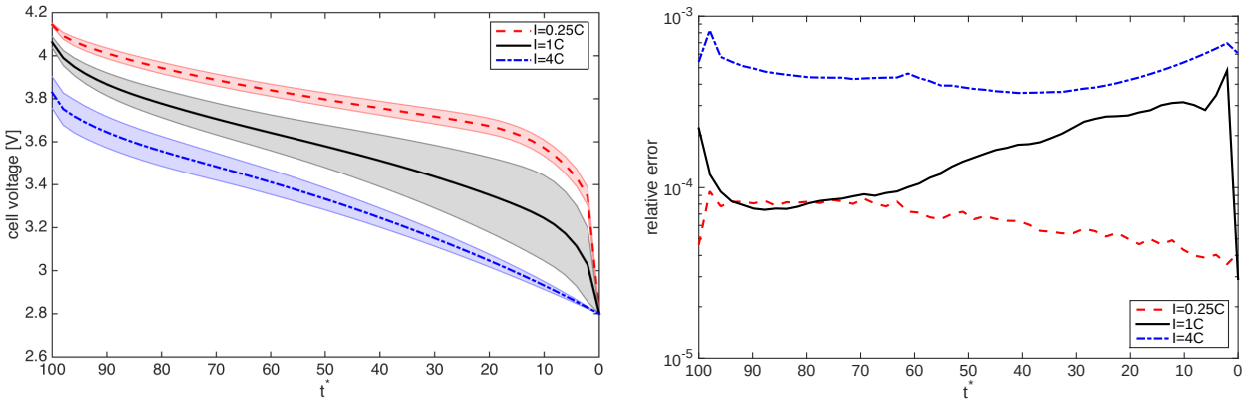
3.6.2 Statistics in normalized time

Since the input samples are different, the battery realizations reach the cut-off potential of 2.8 V at different times. As an example, for 1C rate of discharge, some realizations of the battery reach the cut-off potential before $t = 3200$ s while other realizations need more time, e.g., $t_{max} = 3426$ s. This asynchronous behavior leads to non-smooth dependence of the cell voltage to random inputs and deteriorates the accuracy of PC approximation when the battery approaches the end of discharge.

To overcome this issue, we introduce an uncertain time scaling t^*. In order to rescale the deterministic realizations, we assume at $t = 0$, $t^* = 100$, and we set $t^* = 0$ when the battery reaches the cut-off potential of 2.8 V. Hence, at $t^* = 100$ the battery is fully charged and at $t^* = 0$ it is fully discharged. This rescaling approach enables us to maintain the accuracy of approximation over the entire discharge process without increasing the computational cost or complexity of the problem. We note that our definition of t^* involves a constraint on the cell voltage, such that at $t^* = 0$ the cell voltage for all realizations is equal to 2.8 V, which translates to zero variability for the cell voltage at $t^* = 0$. Our t^* definition imposes no constraints on other QoI such as solid and liquid phase concentrations.

3.6.3 Effects of input uncertainties on cell voltage and concentrations

Quantification of the effects of input variations on the cell voltage and concentrations over the charge or discharge processes could provide a better understanding of the cell behavior. In the following, we present the effects of LIB input uncertainties on the stochastic behavior of the cell voltage ϕ_{cell}, liquid phase concentration c, and solid phase concentration at the surface of the solid particle c_s^{surf} as functions of the normalized time t^*. In order to study the variations of c in all three main regions of the LIB, we compute c at three different locations in the cell; middle of anode, separator and cathode. Similarly, c_s^{surf} is computed in the middle of electrodes.



(a) (b)

Figure 3.4: Effects of input uncertainties on the cell voltage for $I = 0.25$C, 1C and 4C rates of discharge. (a) Mean and standard deviation of the cell voltage. The shaded areas are probability bounds of three standard deviations around the mean; (b) Relative error.

Fig. 4.6(a) shows the mean and probability bounds of three standard deviations around the mean of cell voltage as the functions of t^* for $I = 0.25$C, 1C and 4C. Although at t^* = 100, larger variabilities in ϕ_{cell} correspond to higher discharge rates, ϕ_{cell} experiences its largest standard deviation at 1C rate of discharge, highlighting again the importance of discharge rate on UQ analysis of the LIBs. The validation error of the PC solution constructed from $N = 1000$ samples is displayed in Fig. 4.6(b). In fact, $N = 1000$ samples were enough to accurately approximate c and c_s^{surf} as well.

The corresponding total Sobol' indices of the discharge curves in Fig. 4.6 are presented in Fig. 3.5, from which we highlight that:

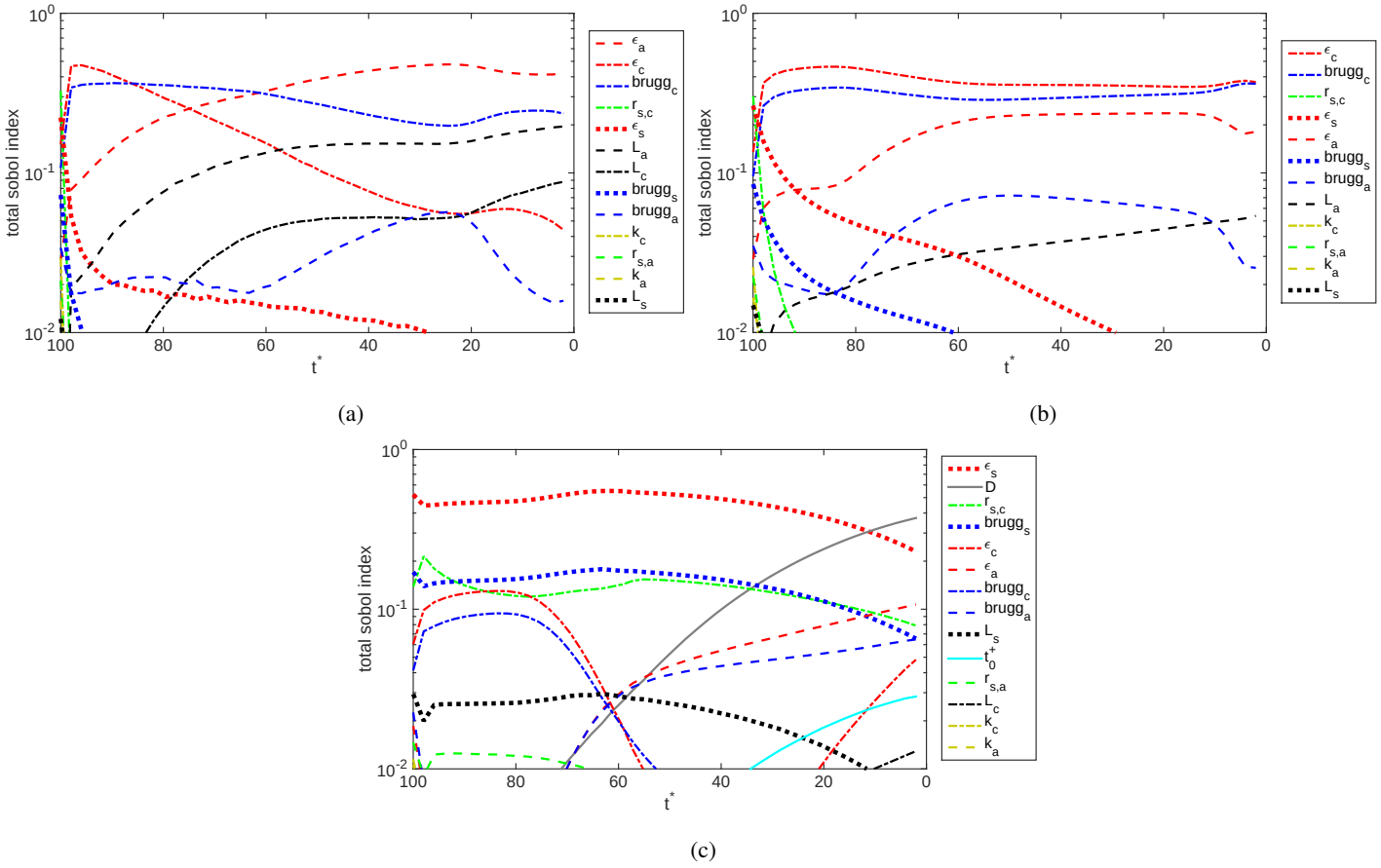


Figure 3.5: Global sensitivity analysis of the cell voltage for: (a) $I = 0.25$C; (b) $I = 1$C; (c) $I = 4$C.

- Similar to the cell capacity, porosity ϵ and Bruggeman coefficient brugg in all three regions are among the most important random inputs which contribute to the variability of the cell voltage for all three discharge rates.

- For all three discharge rates, unlike the cell capacity, uncertainties in k_a, k_c, and $r_{s,a}$ contribute to the cell voltage variations, although, their impacts are limited to the onset of discharge for low to medium discharge rates.

- The uncertainty in σ_a, σ_c, $D_{s,a}$, and $D_{s,c}$ has no significant effects on the variability of the cell voltage.

- As the discharge rate increases, random parameters in separator, i.e., ϵ_s, brugg_s, and L_s, contribute

more and more to the cell voltage variations, such that for 4C rate of discharge, ϵ_s becomes the most important random input affecting the cell voltage variability. A similar behavior was observed in Fig. 3.3.

- For $I = 4$C, variations in the cell voltage are highly affected by the uncertainties in the diffusion coefficient of the liquid phase D. Although, for low to medium discharge rates, D is an insignificant random input, for high discharge rates, its Sobol' index S_D^T grows considerably. Moreover, unlike the cell capacity, Li^+ transference number t_+^0 has a high contribution in the cell voltage variability, when the battery is discharged at 4C rate.

- Similar to the cell capacity, effects of uncertainties in the length of electrodes, L_a and L_c, on the variations of ϕ_{cell} decreases as the rate of discharge increases.

Figs. 3.6(a)-(c) show the mean values of c and their three standard deviation bounds at three locations for 0.25C, 1C and 4C rates of discharge. As the LIB is discharged at higher rates, larger concentration gradients are developed across the cell to balance the migration of anions [90], as a result of which the liquid concentration in the cathode approaches to zero near the end of discharge for 4C discharge rate (Fig. 3.6(c)). This may lead to zero concentrations in a region at the back of the cathode electrode, which means the active materials may not be utilized further. This limits the capability of the LIB to be discharged at higher rates. On the other hand, high concentrations in the anode electrode may be problematic when the lithium salt/solvent system used has a solubility limit. For example, such a limit is 2100 mol/m^3 for Lithium Perchlorate in Propylene Carbonate at room temperature [90]. Consequently, analysis of the variability of liquid phase concentration near the end of discharge may provide a more accurate understanding of the rate limiting mechanisms associated with diffusion of lithium ions. As can be observed from this figure, our definition of t^* does not impose any constraints on c; hence, unlike the cell voltage, variability of c is not equal to zero at $t^* = 0$.

As mentioned in Section 3.5, it has been shown that σ and D_s depend on the solid phase concentration. Taking into account the variability in the solid phase concentration due to input variations may help to assess such dependencies more accurately. A similar analysis is performed for the solid phase concentration at the

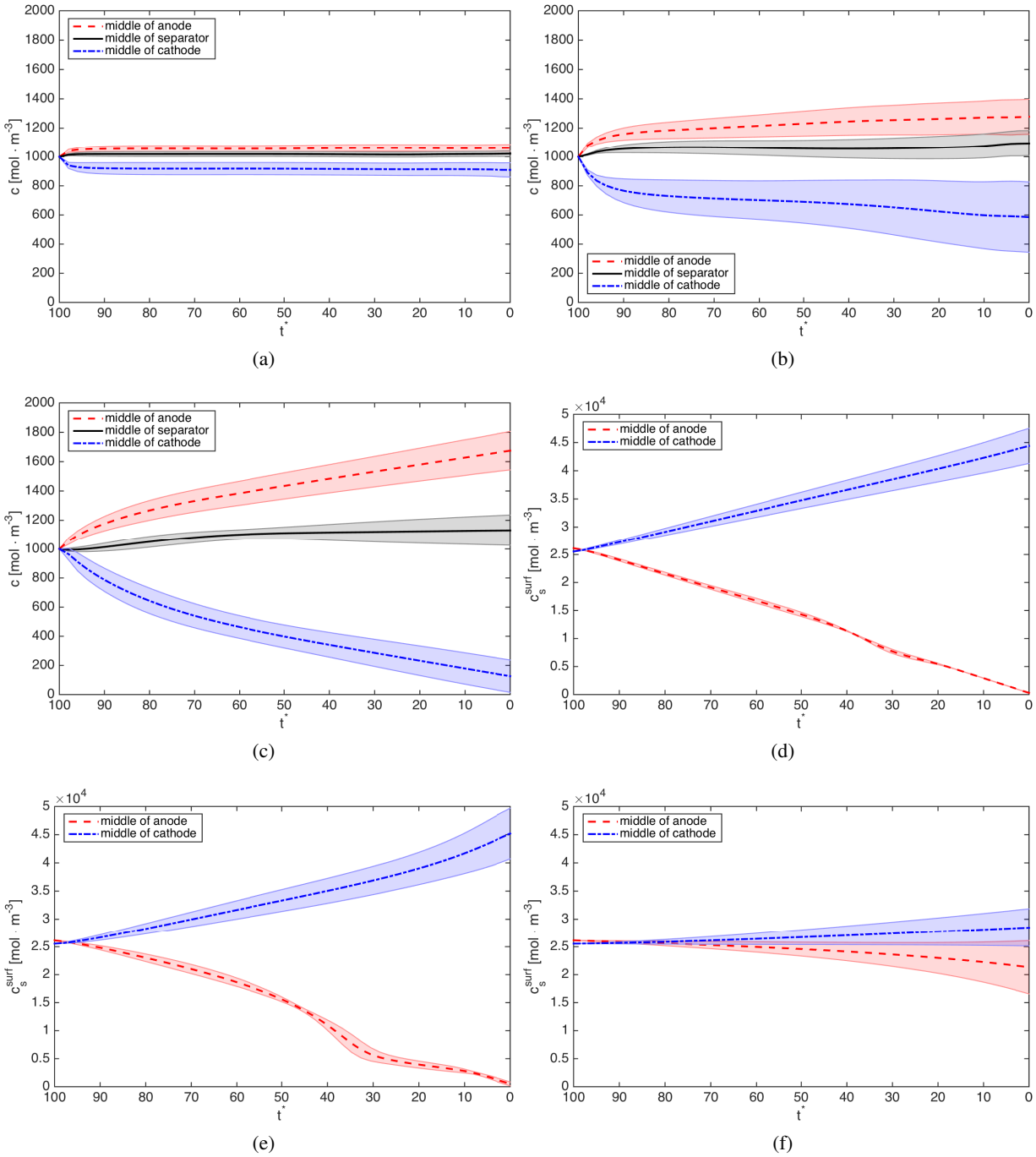


Figure 3.6: Mean and standard deviation of the liquid phase concentration c in the middle of anode, separator and cathode for: (a) $I = 0.25$C; (b) $I = 1$C; (c) 4C rates of discharge. Mean and standard deviation of the solid phase concentration c_s^{surf} in the middle of anode and cathode for: (d) $I = 0.25$C; (e) $I = 1$C; (f) 4C rates of discharge. The shaded areas are probability bounds of three standard deviations around the mean.

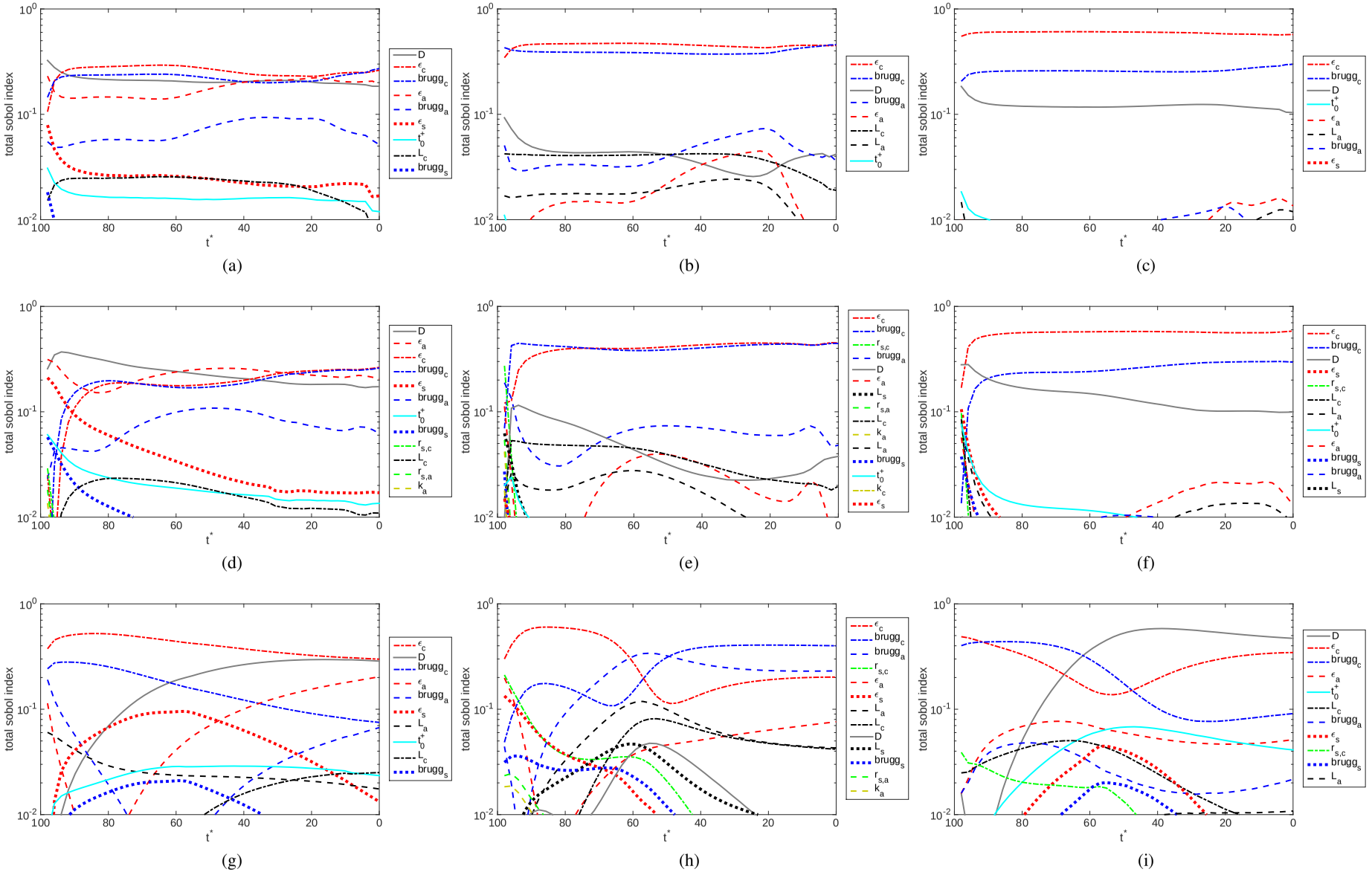


Figure 3.7: Global sensitivity analysis of the liquid phase concentration for $I = 0.25$C, 1C and 4C rates of discharge. (a) c in the middle of anode with $I = 0.25$C; (b) c in the middle of separator with $I = 0.25$C; (c) c in the middle of cathode with $I = 0.25$C; (d) c in the middle of anode with $I = 1$C; (e) c in the middle of separator with $I = 1$C; (f) c in the middle of cathode with $I = 1$C; (g) c in the middle of anode with $I = 4$C; (h) c in the middle of separator with $I = 4$C; (i) c in the middle of cathode with $I = 4$C.

surface of the solid particle c_s^{surf}. The mean values of c_s^{surf} bounded by three standard deviations in the middle of electrodes for $I = 0.25$C, 1C and 4C rates of discharge are also presented in Figs. 3.6(d)-(f). It can be observed that the variability of c_s^{surf} in the middle of the anode does not grow significantly as the battery is discharged at low and medium rates, while for the high discharge rate, standard deviation of c_s^{surf} in both electrodes is increased monotonically over the course of discharge.

Fig. 3.7 shows the corresponding total Sobol' indices of c in three regions for $I = 0.25$C, 1C and 4C rates of discharge. We observe that for low to medium discharge rates, the influential random inputs on the variability of c vary considerably from one region to another. As an instance, for $I = 0.25$C, variations in c in the middle of cathode are mainly affected by uncertainty in ϵ_c, brugg_c, and D, while in the anode, theses variations are affected by the uncertainty in D, ϵ_c, brugg_c, ϵ_a, brugg_a, ϵ_s, t_+^0, L_c, and brugg_s. In other words, for low discharge rates, variability of c in the anode is affected by the random inputs associated with the other two regions, i.e., separator and cathode, while this is not true for the variation of c in the cathode. Additionally, we note that:

- Similar to the cell capacity and voltage, uncertainties in σ_a, σ_c, $D_{s,a}$, and $D_{s,c}$ have no significant effects on the variability of c. Moreover, although reaction rate constants k_a and k_c appear in Figs. 3.7(d), (e), and (h), their impact on the variability of c is small.

- Despite the discharge rate in all three regions, uncertainty in D and t_+^0 has more pronounced effects on the variation of c in both electrodes in comparison to separator.

- S_D^T is increased in the electrodes and decreased in the separator as the LIB is discharged at higher rates.

- Unlike the cell capacity and voltage, the effect of uncertainties in the length of the electrodes L_a and L_c on the variations of c increases as the rate of discharge increases.

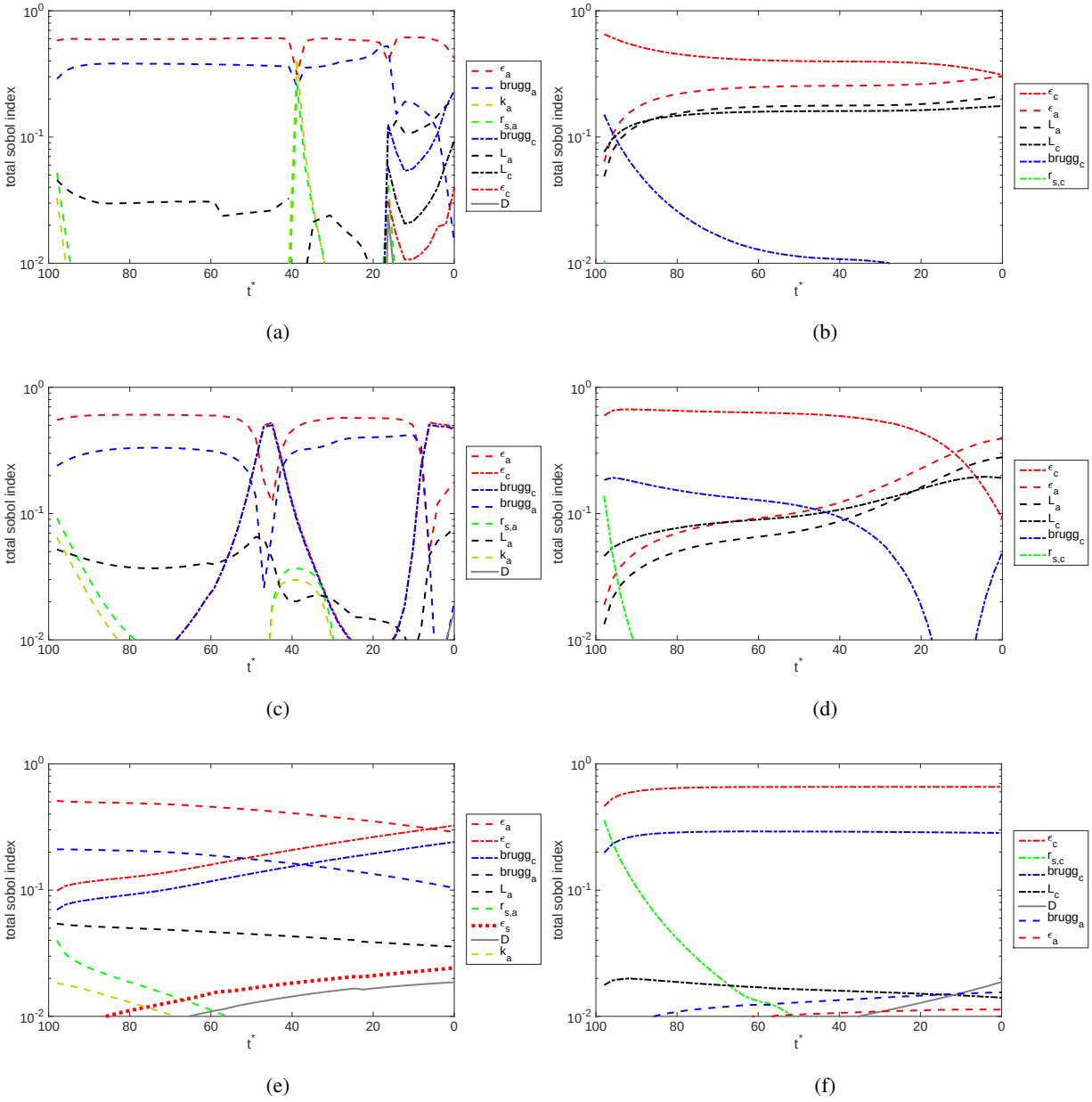


Figure 3.8: Sensitivity analysis of the solid phase concentration for $I = 0.25$C, 1C and 4C rates of discharge. (a) c_s^{surf} in the middle of anode with $I = 0.25$C; (b) c_s^{surf} in the middle of cathode with $I = 0.25$C; (c) c_s^{surf} in the middle of anode with $I = 1$C; (d) c_s^{surf} in the middle of cathode with $I = 1$C; (e) c_s^{surf} in the middle of anode with $I = 4$C; (f) c_s^{surf} in the middle of cathode with $I = 4$C.

A global SA of c_s^{surf} in both electrodes is also presented in Fig. 3.8. We see that uncertainty in the solid phase diffusion coefficient D_s has no significant effects on the variability of cell capacity, voltage and

liquid phase concentration. Surprisingly, D_s may be labeled as an insignificant random input even on the variations of c_s^{surf} in both electrodes for all three discharge rates. Moreover, the uncertainty in σ_a and σ_c also has no impacts on the variability of c_s^{surf} independent of the discharge rate. As the LIB experiences higher discharge rates, the total Sobol' index of $r_{s,a}$ decreases, while for $r_{s,c}$, an increase in the total Sobol' index is observed.

In overall, we see that in our stochastic LIB example, the uncertainty in σ and D_s has no significant effects on the variability of the cell capacity, voltage and concentrations, hence, may be treated as deterministic inputs. Because of this, tight quality control measures are not needed for these parameters. On the other hand, ϵ and brugg are the most important random inputs independent of the discharge rate, and require tight quality control measures to reduce the LIB cell-to-cell variations. Moreover, we observe that the relative contribution of uncertainty in the model parameters in the overall performance variability is highly affected by the battery discharge rate.

We emphasize that the results presented in this section apply only to the $LiC_6/LiCoO_2$ cell we consider in this study subjected to the input variabilities presented in Table 3.2. Any changes in these uncertainties or cell chemistry/configuration may result in different observations. This is also the case if a more accurate LIB model, which accounts for additional physical and chemical phenomena, is employed.

3.7 Summary and Conclusions

A sampling-based UQ approach was introduced to study the effects of input uncertainties on the performance of the LIBs. The proposed UQ approach is based on polynomial chaos expansion framework and hinges on a sparse approximation technique to achieve an accurate estimation of solution statistics with a small number of LIB simulations. Additionally, the proposed method enables one to identify the most important random inputs for any QoI such as capacity, voltage, and concentrations by performing a global sensitivity analysis via computing the total Sobol' indices. Such an analysis is helpful in designing more efficient and targeted quality control measures, from material selection to cell assembly, and reducing the manufacturing cost of the LIBs.

Performance of the proposed UQ approach was explored through its application to an $LiC_6/LiCoO_2$

LIB discharged at three different rates. It was shown that the proposed UQ method can accurately compute the variability in the output QoIs such as the cell capacity, voltage, and concentrations with a small number of battery simulations, 1000 in this example. The global sensitivity analysis results corresponding to these QoIs showed that the identification of the most important uncertain inputs is highly affected by the battery discharge rate. For all three discharge rates, we found that the porosity and Bruggeman coefficient are among the most significant uncertain parameters in the performance variability of the examined LIB.

We acknowledge that the LIB model we considered in this study suffers from many unmodeled phenomena such as side reactions, stresses associated with volume changes, cell degradation, and temperature dependence, which may affect the stochastic behavior of the cell. Since the proposed stochastic LIB model is based on sampling, in order to include these phenomena in the model, more accurate LIB models can be incorporated without changing the overall UQ framework. This may, however, increase the number of required battery simulations and, hence, the overall computational cost. Additionally, the accuracy of the uncertainty models we adopted for the LIB input parameters is limited by the available experimental data. For some of these inputs, data are highly sparse or non-existent. We, therefore, resorted to assumptions in describing the uncertainty. When data becomes available, these uncertainty models may be improved accordingly.

The proposed UQ framework in this thesis was developed for forward propagation of the uncertainties in the LIB simulations. Interesting future research directions include using this forward UQ framework along with actual experimental data to infer battery model parameters as well as quantitative validation of the model itself.

Nomenclature

D	diffusion coefficient of the liquid phase [$\mathrm{m}^{-2} \cdot \mathrm{s}^{-1}$]
D_s	diffusion coefficient of the solid phase [$\mathrm{m}^{-2} \cdot \mathrm{s}^{-1}$]
F	Faraday's constant = 97484 [$\mathrm{C} \cdot \mathrm{mol}^{-1}$]
I	total current density across the stack [$\mathrm{amp} \cdot \mathrm{m}^{-2}$]
L	width [m]
N	number of samples
P	number of PC basis functions of pth total order in dimension d
S_k	first order Sobol' index of kth random input
S_k^T	total Sobol' index of kth random input
T	temperature [K]
V	open circuit potential of the active material [V]
Ω	sample set
$\alpha_{\boldsymbol{i}}$	PC coefficients
$\boldsymbol{\xi}$	vector of input random variables
δ	Kronecker delta or Dirac delta function
ϵ	porosity of electrodes and stack
η	over-potential in electrodes [V]
γ	the truncation error tolerance of BPDN problem
κ	electronic conductivity of the liquid phase [$\mathrm{S} \cdot \mathrm{m}^{-1}$]
κ_D	liquid phase diffusional conductivity [$\mathrm{S} \cdot \mathrm{m}^{-1}$]
brugg	Bruggeman coefficient
$\mathscr{I}_{d,p}$	set of pth order multi-indices in dimension d
ϕ_{cell}	cell voltage [V]
ϕ_e	Li^+ ion potential in liquid phase [V]
ϕ_s	electron potential in the solid phase [V]
$\psi_{\boldsymbol{i}}$	multivariate PC basis functions
$\rho(\xi)$	probability measure of random variable ξ
σ	electronic conductivity of the solid phase [$\mathrm{S} \cdot \mathrm{m}^{-1}$]
τ	tortuosity
a	active particle surface area per unit volume of electrode [$\mathrm{m}^2 \cdot \mathrm{m}^{-3}$]
c	salt concentration in liquid phase [$\mathrm{mol} \cdot \mathrm{m}^{-3}$]
c_s	lithium concentration in solid phase [$\mathrm{mol} \cdot \mathrm{m}^{-3}$]
c_s^{surf}	lithium concentration in solid phase at $r = r_s$ [$\mathrm{mol} \cdot \mathrm{m}^{-3}$]
d	number of random inputs
i_{ex}	exchange current density of an electrode reaction [$\mathrm{amp} \cdot \mathrm{m}^{-2}$]
j_{vol}	volumetric reaction flux in the pore walls [$\mathrm{amp} \cdot \mathrm{m}^{-3}$]
k	reaction rate constant [$\mathrm{m}^4 \cdot \mathrm{mol} \cdot \mathrm{s}$]
p	total order of the PC expansion
r	micro-scale distance from the center of solid particle [m]
r_s	solid particle size [m]

t	time [s]
t_+^0	Li^+ transference number
x	distance from anode [m]

Superscripts

eff	effective value
h	harmonic mean
j	unit flux portion
o	internal mixing portion

Subscripts

$\boldsymbol{i}$	multi-index
$\max$	maximum
a	anode
c	cathode
s	separator

Chapter 4

Optimal Sampling Strategies for Regression-based PCEs

4.1 Introduction

As discussed in Section 1.2.1.2, non-intrusive approximations of the PC coefficients via sampling-based techniques such as least squares regression (LSR) [113], pseudo-spectral collocation [258], Monte Carlo sampling [149], and compressive sampling [59, 105] have been gaining more and more popularity as they do not require modifications of the existing deterministic solvers. In this chapter, we consider the LSR approach to determine the PC coefficients $\boldsymbol{\alpha}$ and review the available sampling techniques for this problem. The LSR technique was discussed in Section 1.2.1.2, but in order to keep this chapter self-contained, we review the LSR approach in the context of PCEs in the following.

Let $(\Omega, \mathcal{T}, \mathcal{P})$ be complete probability space where Ω is the sample set (a.k.a. design space in the context of design of experiments) and $\mathcal{P}$ is a probability measure on the $\sigma-$field $\mathcal{T}$. Also assume that the system input uncertainty has been discretized and approximated by random variables, such that the vector $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_d) : \Omega \to \mathbb{R}^d$, $d \in \mathbb{N}$, represents the set of d independent random inputs with a joint PDF denoted by $\rho(\boldsymbol{\xi})$. The PC representation of the scalar finite variance output QoI denoted by $u(\boldsymbol{\xi})$, in terms of the orthogonal multivariate polynomials $\psi_j(\boldsymbol{\xi})$, then reads

$$u(\boldsymbol{\xi}) = \sum_{j=1}^{\infty} \alpha_j \psi_j(\boldsymbol{\xi}), \tag{4.1}$$

where α_j are the deterministic PC coefficients to be determined. For computational purpose, the infinite series in (4.1) may be truncated by retaining the first P terms as

$$u(\boldsymbol{\xi}) = \sum_{j=1}^{P} c_j \psi_j(\boldsymbol{\xi}) + \epsilon(\boldsymbol{\xi}), \tag{4.2}$$

where $\epsilon(\boldsymbol{\xi})$ represents the random truncation error. A detailed description of this truncation, construction of the orthogonal PC bases $\psi_j(\boldsymbol{\xi})$, and the convergence properties of the PC expansion in (4.2) has been presented in Section 1.2.1.

Let us denote the ith realization of $\boldsymbol{\xi}$ by $\boldsymbol{\Xi}^{(i)}$, generated randomly via Monte Carlo (MC) method, for instance, according to $\rho(\boldsymbol{\xi})$, and the corresponding realization of $u(\boldsymbol{\xi})$ by $u(\boldsymbol{\Xi}^{(i)})$. Given the pair $\{\boldsymbol{\Xi}^{(i)}\}_{i=1}^N$ and $\{u(\boldsymbol{\Xi}^{(i)})\}_{i=1}^N$, with N being the number of independent samples, the discrete representation of (4.2) can be written as

$$\boldsymbol{u} = \boldsymbol{\Psi}\boldsymbol{c} + \boldsymbol{\epsilon}, \tag{4.3}$$

where $\boldsymbol{u} = (u(\boldsymbol{\Xi}^{(1)}), \cdots, u(\boldsymbol{\Xi}^{(N)}))^T \in \mathbb{R}^N$ contains the realizations of the QoI $u(\boldsymbol{\xi})$, $\boldsymbol{\Psi}(i,j) = \psi_j(\boldsymbol{\Xi}^{(i)}) \in \mathbb{R}^{N\times P}$ is the measurement matrix containing samples of the PC basis, and $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_P)^T \in \mathbb{R}^P$ is the vector of unknown PC coefficients, a.k.a. estimators, and the vector $\boldsymbol{\epsilon} \in \mathbb{R}^N$ contains unknown truncation errors between the solution realization $u(\boldsymbol{\Xi})$ and the truncated PC approximation of u. A diagonal positive-definite matrix $\boldsymbol{W}$ is also introduced such that $\boldsymbol{W}(i,i) = w(\boldsymbol{\Xi}^{(i)})$ is a function of the sample points, which indeed depends on the sampling strategy, and will be discussed later in Section 4.4. The unknown coefficients $\boldsymbol{\alpha}$ may be approximated by solving the least squares problem

$$\min_{\boldsymbol{c}} \|\boldsymbol{W}\boldsymbol{u} - \boldsymbol{W}\boldsymbol{\Psi}\boldsymbol{\alpha}\|_2, \tag{4.4}$$

where $\|\cdot\|_2$ is the standard Euclidean norm. When $\boldsymbol{W}\boldsymbol{\Psi}$ is full rank, the solution to (4.4) is computed from the normal equation

$$(\boldsymbol{W}\boldsymbol{\Psi})^T(\boldsymbol{W}\boldsymbol{\Psi})\boldsymbol{\alpha} = (\boldsymbol{W}\boldsymbol{\Psi})^T\boldsymbol{W}\boldsymbol{u}. \tag{4.5}$$

In general, a stable solution $\boldsymbol{\alpha}$ to (4.5) requires $N \geq P$ samples of u [107]. For example, standard sampling techniques such as MC require N to be approximately 3-4 times larger than P. The computational cost of constructing the PCE in (4.2) via (4.4) is controlled by the number of samples N. In complex engineering applications obtaining the samples of the output QoI, u, may be quite expensive. Hence, ideally we would like to reduce the number of samples N required for a stable and accurate solution recovery in (4.4) to reduce the construction cost of the surrogate PCE model of $u(\boldsymbol{\xi})$ in (4.2). Moreover, as we saw in Section 1.2.1, the number of unknown PC coefficients, P, grows quickly as either the stochastic

dimensionality of the problem, d, or the total order of the PCE, p, is increased. Consequently, accurate approximation of $\boldsymbol{c}$ with a limited computational budget N for complex engineering systems which either involve a large number of random inputs, i.e., $d \gg 1$, or require high-order PCEs becomes a critical task.

This has motivated researchers to explore different sampling strategies, beyond the random MC sampling, for constructing the sample set $\{\boldsymbol{\Xi}^{(i)}\}_{i=1}^{N}$ to achieve certain *optimality*, while maintaining LSR solution stability and accuracy. In general, one may categorize the sampling techniques for the LSR problem into two groups: random and deterministic. Unlike random sampling, deterministic sampling techniques attempt to select the sample points $\{\boldsymbol{\Xi}^{(i)}\}_{i=1}^{N}$ in a deterministic fashion.

Among the early works on sampling techniques for regression-based PCEs, we mention [42] in which Latin hypercube (LH) sampling is used instead of the standard MC sampling to improve the accuracy of PC approximation. Recently, a random sampling technique referred to as *asymptotic* sampling was studied in [107] based on asymptotic analysis of d-dimensional Legendre and Hermite polynomials. This approach leads to sampling from a d-dimensional Chebyshev distribution for the case of Legendre polynomials, while sampling uniformly from a d-dimensional ball with a radius determined by the total order of the PC expansion is advised for Hermite polynomials. Additionally, an importance sampling approach denoted by *coherence-optimal* sampling was proposed in [107] which ensures an accurate computation of $\boldsymbol{\alpha}$ with a number of solution realizations that depends linearly on P (up to a logarithmic factor). Random sampling from tensor product Gaussian quadrature points has been also proposed in [271], where the authors prove an asymptotic (in PC order) stability of this method with a number of points which scales linearly (up to a logarithmic factor) with P.

A quasi-Monte Carlo (QMC) method using the so-called low-discrepancy points was employed in [159] to deterministically select the sample points for the regression-based PCE such that a stable and accurate solution is achieved with a number of samples N proportional to the square of the size of PC basis, P, up to logarithmic factors. Another deterministic sampling approach based on Weil's theorem was proposed in [272] which leads to a stable approximation in the Chebyshev basis with a number of samples that scales quadratically in P.

In the context of design of experiments (DOE), optimal sampling of the design points $\{\boldsymbol{\Xi}^{(i)}\}_{i=1}^{N}$ has

received extensive attention in the statistics community and for many science and engineering applications [25, 9, 84, 189, 83]. The alphabetic optimality criteria for DOE, a.k.a. *classical optimality*, either focus on minimizing the error in estimating the unknown coefficients (a.k.a. parameters) $\boldsymbol{\alpha}$, e.g., *D*-, *A*-, and *E*-optimal designs, or the error in the model prediction u, e.g., *I*-optimal design. These optimality criterion are mostly based on some function of the information matrix $\boldsymbol{M}_{\boldsymbol{\Psi}}$ given by

$$\boldsymbol{M}_{\boldsymbol{\Psi}} = \boldsymbol{\Psi}^T \boldsymbol{\Psi}, \tag{4.6}$$

when the standard LSR is considered, and

$$\boldsymbol{M}_{\boldsymbol{\Psi}} = (\boldsymbol{W}\boldsymbol{\Psi})^T (\boldsymbol{W}\boldsymbol{\Psi}), \tag{4.7}$$

for the case of weighted LSR problem (4.4). For example, *D*-optimal criterion takes the determinant of $\boldsymbol{M}_{\boldsymbol{\Psi}}^{-1}$, denoted by $|\boldsymbol{M}_{\boldsymbol{\Psi}}^{-1}|$, as an overall measure of the variance of the estimated regression coefficients and constructs $\{\boldsymbol{\Xi}^{(i)}\}_{i=1}^{N}$ such that $|\boldsymbol{M}_{\boldsymbol{\Psi}}^{-1}|$ is minimized. Among the limited works in the literature, which has utilized the alphabetic optimality criteria in the context of regression-based PCE, we mention [268] in which *D*-optimality criterion is used to generate the sample set. It was shown that a Fedorov-genetic algorithm used to generate the *D*-optimal sample set could compute the PC coefficients accurately with a smaller cost in comparison to other common sampling techniques. In [30], D-optimal criterion has been employed to efficiently estimate the PCE-based Sobol' indices via the LSR approach. Very recently, a quasi-optimal sampling approach for the weighted LSR problem has been proposed in [219] which takes a function that involves $|\boldsymbol{M}_{\boldsymbol{\Psi}}|$ as the optimality measure.

As the literature review indicates, there are several recent studies that are dedicated to explore more advanced sampling techniques to be used in the context of regression-based PCEs. In order to address this growing interest among the researchers, we aim to report the most recent advancements in this area, as well as providing a comprehensive review of the most popular sampling techniques available in the literature. In addition, we propose *alphabetic-coherence-optimal* sampling techniques: hybrid methods which combine the coherence-optimal approach of [107] with alphabetic optimality criteria. Our ultimate goal is to provide practitioners with a collection of available options and a set of comparative studies as how these various sampling techniques perform.

4.2 Problem statement

The uncertain inputs $\boldsymbol{\xi}$ propagate through the physical system defined on a bounded domain $\mathcal{D} \subset \mathbb{R}^D$, $D \in \{1, 2, 3\}$, with boundaries denoted by $\partial\mathcal{D}$, and governed by

$$\begin{aligned} \mathcal{L}(\boldsymbol{x}, t, \boldsymbol{\xi}; u(\boldsymbol{x}, t, \boldsymbol{\xi})) &= 0, \quad (\boldsymbol{x}, t) \in \mathcal{D} \times [0, T], \\ \mathcal{B}(\boldsymbol{x}, t, \boldsymbol{\xi}; u(\boldsymbol{x}, t, \boldsymbol{\xi})) &= 0, \quad (\boldsymbol{x}, t) \in \partial\mathcal{D} \times [0, T], \\ \mathcal{I}(\boldsymbol{x}, 0, \boldsymbol{\xi}; u(\boldsymbol{x}, 0, \boldsymbol{\xi})) &= 0, \quad (\boldsymbol{x}, t) \in \mathcal{D} \times [0, T], \end{aligned} \tag{4.8}$$

where the differential operator $\mathcal{L}$, initial conditions $\mathcal{I}$, and boundary conditions $\mathcal{B}$ depend on the physics of the problem and $u(\boldsymbol{x}, t, \boldsymbol{\xi}) : \bar{\mathcal{D}} \times [0, T] \times \Omega \to \mathbb{R}$ is the solution to (4.8). For the sake of a clear and simple notation, we suppress the dependence of u on $\boldsymbol{x}$ and t in the rest of this chapter.

4.3 Definitions and background

In Section 4.4, we provide a review of the the sampling techniques mentioned in Section 4.1. In order to keep this chapter self-contained, we next review some important definitions which will be used in the presentation of sampling strategies.

4.3.1 Coherence parameter definition

Let $B^2(\boldsymbol{\Xi})$ represent a uniformly least upper bound on the sum of the squares of the PC basis functions defined as

$$B(\boldsymbol{\Xi}) := \sqrt{\sum_{j=1}^{P} |\psi_j(\boldsymbol{\Xi})|^2}. \tag{4.9}$$

The idea behind this definition is the fact that bounding the spectral radius of the weighted measurement matrix $\boldsymbol{W}\boldsymbol{\Psi}$ yields a bound on the number of samples N necessary to compute the PC coefficients via the LSR problem (4.4). This function has been used in [107, 45] (similarly in [108, 35] in the context of compressive sampling via l_1-minimization) to generate a condition-type number in order to guarantee a stable solution to (4.4).

We also let $G(\boldsymbol{\Xi})$ represent an upper bound on $B(\boldsymbol{\Xi})$ for all $\boldsymbol{\Xi} \in \Omega$,

$$G(\boldsymbol{\Xi}) \geq B(\boldsymbol{\Xi}), \qquad \forall\, \boldsymbol{\Xi} \in \Omega. \tag{4.10}$$

We consider a set of random variables $\boldsymbol{Y}$ as a copy of $\boldsymbol{\xi}$ with a joint PDF given by

$$\rho_{\boldsymbol{Y}}(\boldsymbol{\Xi}) := c^2 \rho(\boldsymbol{\Xi}) G^2(\boldsymbol{\Xi}), \tag{4.11}$$

with the normalizing constant c

$$c = \left(\int_{\Omega} \rho(\boldsymbol{\Xi}) G^2(\boldsymbol{\Xi}) d\boldsymbol{\Xi} \right)^{-1/2}. \tag{4.12}$$

We note that the set $\{\psi_j(\boldsymbol{Y})\}_{j=1}^{P}$ is no longer (approximately) orthogonal under $\rho_{\boldsymbol{Y}}$. In other words, if we sample the random inputs according to $\rho_{\boldsymbol{Y}}$, the PC bases are no longer orthonormal. Following the discussion in [107, Section 2.1], the weight functions, i.e., $\boldsymbol{W}(i,i) = w(\boldsymbol{\Xi}^{(i)})$ in (4.4), with $\boldsymbol{\Xi}^{(i)}$ being the ith realization of $\boldsymbol{Y}$,

$$w(\boldsymbol{Y}) = \frac{1}{cG(\boldsymbol{Y})}, \tag{4.13}$$

are considered to provide (approximate) orthogonality to the set $\{w(\boldsymbol{Y})\psi_j(\boldsymbol{Y})\}_{j=1}^{P}$.

In [107], the *coherence* parameter $\mu(\boldsymbol{Y})$ given by

$$\mu(\boldsymbol{Y}) := \sup_{\boldsymbol{\Xi} \in \Omega} \sum_{j=1}^{P} |w(\boldsymbol{\Xi})\psi_j(\boldsymbol{\Xi})|^2, \tag{4.14}$$

is used to bound the realized spectral radius of $\boldsymbol{W}\boldsymbol{\Psi}$, hence, bounding the number of required samples N, for bounded polynomials, e.g., Legendre polynomials. For the cases in which $\mu(\boldsymbol{Y})$ given by (4.14) is infinite, i.e., for unbounded polynomials such as Hermite polynomials, a coherence parameter defined on an appropriately truncated subset $\mathcal{S}$ of Ω is considered instead. The interested reader is referred to [107, Section 2.2] for further details.

Notation 1. *In order to have a clear and simple notation, we refer to all realized random vectors by* $\boldsymbol{\Xi}$ *regardless of the sampling distribution used to generate these realizations. We also note that the weight functions depend on the distribution used to generate* $\boldsymbol{\Xi}$*.*

4.3.2 Star discrepancy definition

Star discrepancy is a measure of the uniformity of the sample points and is used in the context of low-discrepancy sequences for quasi-Monte Carlo sampling techniques discussed in Section 4.4.6.1. Given

the sample set $\{\boldsymbol{\Xi}^{(i)}\}_{i=1}^{N}$ in d-dimensional unit cube $[0,1)^d$, the star discrepancy D_N^* of the set $\{\boldsymbol{\Xi}^{(i)}\}_{i=1}^{N}$ is defined as

$$D_N^* := \sup_{J\in\mathcal{J}^*} \left| \frac{1}{N} \cdot \#\{i : \boldsymbol{\Xi}^{(i)} \in J, 1 \leq i \leq N\} - \lambda_d(J) \right|, \tag{4.15}$$

where $\mathcal{J}^*$ is the class of all subintervals J of $[0,1)^d$ given by

$$J = \prod_{k=1}^{d} [0, v_k)^d, \qquad 0 \leq v_k \leq 1, \tag{4.16}$$

$\#\mathcal{A}$ denotes the number of elements of the set $\mathcal{A}$, $\lambda_d(J)$ is the Lebesgue measure of J [71]. A well uniformly distributed set of points has a small star discrepancy [32, 173].

4.4 Sampling Techniques

In this section, we start reviewing the standard MC technique followed by the asymptotic and coherence-optimal sampling methods presented in Sections 4.4.3 and 4.4.2, respectively. We then continue by stochastic collocation via quadrature points and randomized quadrature samplings discussed in Section 4.4.4. In Section 4.4.5, we provide a detailed description of the classical optimal designs and lay out a sequential technique to construct these designs. Finally, space-filling designs such as QMC and LH sampling techniques which are popular in designing computer experiments are presented in Section 4.4.6.

4.4.1 Standard Monte Carlo sampling

The natural method of constructing the sample set $\{\boldsymbol{\Xi}^{(i)}\}_{i=1}^{N}$, known as the standard MC, is by sampling $\boldsymbol{\xi}$ according to the orthogonality measure $\rho(\boldsymbol{\Xi})$ of the PC basis, for which the weights are $w(\boldsymbol{\Xi}) = 1$. For instance, standard MC sampling of the d-dimensional Legendre polynomials corresponds to sampling from the uniform distribution on $[-1,1]^d$. For other polynomials, the corresponding sampling distributions are listed in Table 1.1.

As shown in [107, Theorem 3.1], a MC sampling of the d-dimensional Legendre polynomials of total order p gives a coherence parameter of $\mu(\boldsymbol{\xi}) \leq \exp(2p)$, while μ associated with MC sampling of the d-dimensional Hermite polynomials is bounded by $\mu(\boldsymbol{\xi}) \leq C_p \eta_p^p$, with C_p being a constant and η_p a function of p. For high-order PCE's, μ is large and therefore N needs to be large. This is why, for those

cases, alternative sampling distributions are needed to reduce the number of required samples. Convergence analyses of the standard MC method for univariate and multivariate Legendre polynomials are also presented in [160] and [161], respectively.

4.4.2 Coherence-optimal sampling

An important parameter that controls the stability of the LSR problem in (4.4) is the spectral radius of $\boldsymbol{W}\boldsymbol{\Psi}$ denoted by $\rho(\boldsymbol{W}\boldsymbol{\Psi})$. In [107, Lemma 5.1], it is shown that the smaller $\rho(\boldsymbol{W}\boldsymbol{\Psi})$ means more stable and accurate solution to (4.4) could be obtained with a smaller number of samples N. This has motivated the work in [107] to bound the number of samples required for an accurate and stable estimation of the PC coefficients via the LSR problem by bounding $\rho(\boldsymbol{W}\boldsymbol{\Psi})$. The coherence parameter μ defined in (4.14) has been used in this work as a measure to bound $\rho(\boldsymbol{W}\boldsymbol{\Psi})$, and in [107, Theorem 2.1] it is shown that reducing μ reduces the number of required samples for a stable solution recovery in (4.4).

In order to minimize the coherence parameter μ in [107], samples of the inputs are drawn from an alternative measure given by (4.11), i.e., $f_{\boldsymbol{Y}}(\boldsymbol{\Xi})$, instead of sampling from the orthogonality measure $f(\boldsymbol{\Xi})$. For arbitrary values of p and d, an optimal sampling approach, denoted by *coherence-optimal* sampling, that takes $G(\boldsymbol{\Xi})$ in (4.11) to be $B(\boldsymbol{\Xi})$ given by (4.9) is proposed in [107]. Hence, the coherence-optimal approach generates samples according to the measure $f_{\boldsymbol{Y}}(\boldsymbol{\Xi})$ given by

$$f_{\boldsymbol{Y}}(\boldsymbol{\Xi}) := c^2 f(\boldsymbol{\Xi}) B^2(\boldsymbol{\Xi}), \tag{4.17}$$

with the weight function $w(\boldsymbol{\Xi}) = 1/B(\boldsymbol{\Xi})$. An analytic expression for $B(\boldsymbol{\Xi})$ is not generally available; however, for a given set of PC basis, $B(\boldsymbol{\Xi})$ can be evaluated at any arbitrary values of $\boldsymbol{\xi}$. This, and to avoid computing the normalization constant c, has motivated using a Markov Chain Monte Carlo approach to sample from $f_{\boldsymbol{Y}}(\boldsymbol{\Xi})$ in (4.17). A Detailed description of this MCMC technique can be found in [107, Section 4.3.1] and [108, Algorithm 1].

Following [107], the coherence-optimal sampling approach ensures a stable computation of $\boldsymbol{\alpha}$ with a number of solution realizations that depends linearly (up to a logarithmic factor in P) on the number of PC coefficients P. Moreover, through the application of this approach to various numerical examples, it

was empirically observed that the coherence-optimal sampling leads to either similar or considerably more accurate results in comparison to asymptotic and standard MC sampling techniques [107].

4.4.3 Asymptotic sampling

In the asymptotic regime (for p) $G(\boldsymbol{\Xi})$ is taken to be an asymptotic envelope for the polynomials. This approach denoted by the *asymptotic* sampling in [107] leads to sampling from a distribution with i.i.d. Chebyshev components over $[-1, 1]$, i.e.,

$$\rho_Y(\Xi) := \frac{1}{\pi\sqrt{1-\Xi^2}}, \tag{4.18}$$

with the corresponding weights given by

$$w(\boldsymbol{\Xi}) := \prod_{k=1}^{d} (1-\Xi_k^2)^{1/4}. \tag{4.19}$$

In [107, Theorem 3.2] it is shown that a coherence parameter of $\mu(\boldsymbol{Y}) \leq 3^d$, regardless of the relationship between d and p, holds when one samples from the Legendre polynomials according to (4.18) and (4.19). This coherence parameter is independent of the total order of the approximation p, hence, making the Chebyshev sampling suitable for the cases with $p > d$.

An asymptotic analysis of the Hermite polynomials which utilizes Hermite Functions suggests sampling uniformly from a d-dimensional ball of radius $\sqrt{2}\sqrt{2p+1}$, with weights given by $w(\boldsymbol{\Xi}) := \exp(-\|\boldsymbol{\Xi}\|_2^2/4)$ [107, Section 3.2]. The corresponding coherence parameter is given by $\mu(\boldsymbol{Y}) = \mathcal{O}((2p)^{d/2}/\Gamma(d/2+1))$, with Γ being the Gamma function, which demonstrates a weaker dependence on p in comparison to d. We refer the interested reader to [107, Section 3.2] for a detailed description of the uniform sampling from a d-dimensional ball.

4.4.4 Stochastic collocation via quadrature points

A widely used sampling-based approach to compute the PC coefficients is the stochastic collocation (SC) method [153, 259, 11, 47]. The main idea behind the SC technique is to sample the output QoI at particular points, e.g., Gaussian quadrature points, in the stochastic space and then approximate the solution

via interpolation or the solution statistics by numerical integration. SC may also be used to compute the PC coefficients via the projection equation $\alpha_{\boldsymbol{i}} = \mathbb{E}[u\psi_{\boldsymbol{i}}]/\mathbb{E}[\psi_{\boldsymbol{i}}^2]$. In particular, the numerator in this equation is a d-dimensional integral over Ω which may be computed using tensor product or sparse grids constructed from, for instance, one-dimensional Gaussian quadrature points [46].

Although it has been shown that SC based on full tensor products is an effective technique for low-dimensional stochastic problems [12], but it suffers from the so-called *curse-of-dimensionality* as the stochastic dimensionality d is increased: the number of collocation points N in the full tensor products grows exponentially fast as a function of d, i.e., $N = (p+1)^d$, where p is the total order of PCE in each direction [259]. For high-dimensional problems, sparse tensor product spaces first introduced by Smolyak [226] has been proposed to alleviate this issue associated with full tensor products [175].

In an attempt to relax the high computational cost of full tensor products in the context of regression-based PCEs, Zhou et al. [271] proposed to randomly select a subset of samples from the full tensor product based on Gaussian quadrature points with Gaussian quadrature weight functions $w(\boldsymbol{\Xi})$. It was shown that this sampling approach, denoted by *randomized quadratures*, is asymptotically (in order) stable with a number of samples which scales linearly (up to a logarithmic factor) with P. The interested reader is referred to [271] for further details on stability properties and convergence analysis of this method. Very recently, a sampling technique based on deterministic selection of samples from full tensor grids is also proposed in [212] as an alternative to randomized quadratures approach.

4.4.5 Optimal design of experiments

Planning an experimental procedure in order to determine the relation between the factors (a.k.a. random inputs), either discrete or continuous, affecting a process and the output of that process prior to performing the actual experiment is studied under the subject of *design of experiments* (DOE). Conducting experiments could be very expensive and time-consuming, hence, it is of interest to extract as much as information possible from a given amount of experimental effort which is addressed in the context of *optimal design of experiments* (ODE). One of the earliest works on ODE was presented in 1918 by Smith [224], but the major contributions took place a few decades later when Kiefer and co-workers explained the theory

behind ODE and proposed a framework for practical implementation of the D-optimal designs [28].

In this work we are interested in model-based DOE where one approximates a complex system with low-order polynomials. Similar to regression-based PCE, LSR is used to estimate the unknown coefficients of the polynomial models in this context. Model-based ODE seeks selecting the realizations of the random inputs such that some optimal properties are provided for the LSR problem. In the following we will review a major class of model-based ODE known as alphabetic optimal designs.

4.4.5.1 Classical optimality criteria

In (4.2), the truncation error ϵ is a deterministic, but unknown parameter for each realization of the random input parameters $\boldsymbol{\Xi}$. In a statistical context, one may interpret the unknown deterministic error ϵ as a random noise ε that has zero mean and an unknown variance σ_ε^2. Consequently, a discrete representation of (4.2) can be written as

$$\boldsymbol{u} = \boldsymbol{\Psi}\boldsymbol{\alpha} + \boldsymbol{\varepsilon}, \tag{4.20}$$

where the vector $\boldsymbol{\varepsilon} \in \mathbb{R}^N$ has elements with zero mean and an unknown variance σ_ε^2. We also assume that the errors in $\boldsymbol{\varepsilon}$ are uncorrelated, meaning that the off-diagonal elements of the covariance matrix of the random noise $\boldsymbol{\varepsilon}$ are all equal to zero and all its diagonal elements are equal to σ_ε^2. In addition, we assume that ε does not depend on the choice of $\boldsymbol{\Xi}$. With these assumptions, it can be shown that the covariance matrix of the estimated coefficients $\mathrm{var}(\boldsymbol{\alpha})$ obtained via the LSR solution of (4.20) is proportional to $\sigma_\varepsilon^2 \boldsymbol{M}_{\boldsymbol{\Psi}}^{-1}$ [83, Page 9], where the information matrix $\boldsymbol{M}_{\boldsymbol{\Psi}}$ is given in (4.6) for the LSR and in (4.7) for the weighted LSR problems.

The diagonal elements of $\mathrm{var}(\boldsymbol{\alpha})$ are the variances of the estimated coefficients $\boldsymbol{\alpha}$, while the off-diagonal elements are the covariances between these estimators. In model-based ODE, it is of interest to construct $\{\boldsymbol{\Xi}^{(i)}\}_{i=1}^N$ such that the corresponding $\mathrm{var}(\boldsymbol{\alpha})$ is the smallest among the all possible choices of constructing the input samples. Consequently, several real-valued functionals of the information matrix, i.e., $\phi(\boldsymbol{M}_{\boldsymbol{\Psi}})$, known as alphabetic or classical optimality criteria in the literature, have been suggested as the measures of *smallness* of the covariance matrix. Alphabetic optimal designs deal with the selection of sample points $\{\boldsymbol{\Xi}^{(i)}\}_{i=1}^N$ from the design space Ω such that the resulting optimality criterion $\phi(\boldsymbol{M}_{\boldsymbol{\Psi}})$ has

the smallest value among the all possible choices of the sample points [165]. Among the most popular alphabetic optimality criteria we mention:

- D-optimality: An optimal design obtained by minimizing the determinant of the inverse of information matrix, i.e., $\phi_D = |\boldsymbol{M}_{\boldsymbol{\Psi}}^{-1}|^{1/P}$. This can be interpreted as minimizing the geometric mean of the errors in the estimated coefficients $\boldsymbol{\alpha}$ [73].

- A-optimality: Takes the trace of the inverse of the information matrix as the overall measure of the average variance of the estimators. That is, $\phi_A = \text{Tr}(\boldsymbol{M}_{\boldsymbol{\Psi}}^{-1})$ is minimized, where $\text{Tr}(\cdot)$ denotes the trace operator. Stated differently, A-optimality measure minimizes the arithmetic mean of the errors in the estimator.

- E-optimality: The largest eigenvalue of $\text{var}(\boldsymbol{c})$ corresponds to the largest error of the estimator. Consequently, the E-optimality criterion minimizes the largest eigenvalue $\lambda_{\min}$ of the inverse of the information matrix, i.e., $\phi_E = \lambda_{\min}(\boldsymbol{M}_{\boldsymbol{\Psi}}^{-1})$ is minimized.

- I-optimality: An optimal design obtained by minimizing the normalized average variance of the prediction $u(\boldsymbol{\Xi})$ over the design space, i.e., $\phi_I = \int_\Omega \boldsymbol{\psi}^T(\boldsymbol{\Xi})\boldsymbol{M}_{\boldsymbol{\Psi}}^{-1}\boldsymbol{\psi}(\boldsymbol{\Xi})f(\boldsymbol{\Xi})d\boldsymbol{\Xi}$, with $\boldsymbol{\psi}(\boldsymbol{\Xi})$ being a realization of the PC basis $\{\psi_j\}_{j=1}^P$.

Remark 6. *We note that the measurement matrix $\boldsymbol{\Psi}$, hence the information matrix $\boldsymbol{M}_{\boldsymbol{\Psi}}$, only depend on the input sample set $\{\boldsymbol{\Xi}^{(i)}\}_{i=1}^N$, and not the realization of the output QoI u.*

Remark 7. *We note that $\text{var}(\boldsymbol{\alpha})$ is equal to $\sigma_\varepsilon^2\boldsymbol{M}_{\boldsymbol{\Psi}}^{-1}$. In fact, since σ_ε^2 is an unknown proportionality constant, it is common in the context of ODE to ignore the term σ_ε^2, i.e., setting $\sigma_\varepsilon^2 = 1$ [101].*

Remark 8. *One can show that the I-optimal criterion simplifies to $\phi_I = Tr(\boldsymbol{M}\boldsymbol{M}_{\boldsymbol{\Psi}}^{-1})$, where the moment matrix $\boldsymbol{M}$ is given by $\boldsymbol{M} = \int_\Omega \boldsymbol{\psi}^T(\boldsymbol{\Xi})\boldsymbol{\psi}(\boldsymbol{\Xi})f(\boldsymbol{\Xi})d\boldsymbol{\Xi}$ (see for example [109, Page 5] or [166, Page 473]). Due to the orthogonality of the PC basis functions, the moment matrix $\boldsymbol{M}$ is an identity matrix of size P and, consequently, the I-optimal criterion reduces to A-optimal criterion, i.e., $\phi_I = Tr(\boldsymbol{M}_{\boldsymbol{\Psi}}^{-1})$, when one approximates the output via PCE.*

Remark 9. *Instead of minimizing ϕ_D, ϕ_A and ϕ_E, one may alternatively maximize $|\boldsymbol{M}_{\boldsymbol{\Psi}}|^{1/P}$, Tr($\boldsymbol{M}_{\boldsymbol{\Psi}}$), $\lambda_{\max}(\boldsymbol{M}_{\boldsymbol{\Psi}})$ to obtain D-, A- and E-optimal designs, respectively [5, 83].*

As mentioned in Section 1, *D*-, *A*-, and *E*-optimal designs are referred to as estimation-oriented optimal designs [123], meaning that these designs are focused on precise estimation of the unknown coefficients $\boldsymbol{\alpha}$. On the other hand, prediction-oriented optimal designs such as I-optimal, a.k.a., *Q*-optimal [84], or *V*-optimal [5] are developed to reduce the variance of the prediction u. I-optimal designs have received increasing attention in the context of response surface designs, where the ultimate goal is to make predictions [123]. There are more optimality criteria that are labeled by other letters in the literature and have been introduced over the time. Discussion about these optimal measures is beyond the scope of the present work and we refer the interested reader to [9, 189, 83, 132, 123] for more details on alphabetic optimal designs. Next, we review the available techniques in the literature to construct *D*-, *A*-, and *E*-optimal designs.

4.4.5.2 Constructing optimal designs

In Section 4.1, we categorized ODE under deterministic sampling techniques for LSR problems. This is in fact true only for limited cases of simple low-order polynomial models with a very limited number of discrete random inputs (factors) where analytic expressions for the optimal designs are available, see, e.g., [118, 130, 53]. In general, however, the ODE optimization problems are solved numerically, often, involving search through a large, but finite, number of randomly generated candidate designs. The selected optimal design, therefore, depends on the choice of the candidate ensembles. This suggests that one may consider ODE as (quasi-)random sampling techniques.

Heuristic exchange algorithms are among the earliest search methods proposed to construct optimal designs [151, 227]. Exchange algorithms were originally developed for D-optimal designs since due to the available simple update equation for the determinant of the information matrix (see Appendix C), they are computationally more feasible in comparison to other criteria [83]. Moreover, it has been shown that generally speaking, D-optimal designs perform well with respect to other criteria [9].

The original Fedorov exchange algorithm (FEA) proposed in [84] starts by generating a discrete candidate set which includes $N_c > N$ realizations of the input random variables, generated based on the

orthogonality measure, for example, and the corresponding $N_c \times P$ candidate measurement matrix $\boldsymbol{\Psi}_c$. Next, a non-singular initial design $\boldsymbol{\Psi}$ of size $N \geq P$ is selected randomly from $\boldsymbol{\Psi}_c$ and FEA iteratively modifies the current design by exchanging a row $\boldsymbol{\psi}_i$ in $\boldsymbol{\Psi}$ with a row $\boldsymbol{\psi}_j$ in $\boldsymbol{\Psi}_c$ which corresponds to the maximum reduction in the D-optimal measure $\phi_D = |\boldsymbol{M}_{\boldsymbol{\Psi}}^{-1}|^{1/P}$. Let $\boldsymbol{\Psi}_{\text{old}}$ and $\boldsymbol{\Psi}_{\text{new}}$ denote the measurement matrix $\boldsymbol{\Psi}$ before and after the exchange, respectively. Using (C.1) the update formula for the determinant is

$$\frac{|\boldsymbol{\Psi}_{\text{new}}^T \boldsymbol{\Psi}_{\text{new}}|}{|\boldsymbol{\Psi}_{\text{old}}^T \boldsymbol{\Psi}_{\text{old}}|} = 1 + \Delta(\boldsymbol{\psi}_i, \boldsymbol{\psi}_j), \tag{4.21}$$

where the Fedorov's Δ-function is given by

$$\begin{aligned} \Delta(\boldsymbol{\psi}_i, \boldsymbol{\psi}_j) &= d(\boldsymbol{\psi}_j) - \left[d(\boldsymbol{\psi}_i)d(\boldsymbol{\psi}_j) - (d(\boldsymbol{\psi}_i, \boldsymbol{\psi}_j))^2\right] - d(\boldsymbol{\psi}_i), \\ d(\boldsymbol{\psi}_i) &= \boldsymbol{\psi}_i^T (\boldsymbol{\Psi}_{\text{old}}^T \boldsymbol{\Psi}_{\text{old}})^{-1} \boldsymbol{\psi}_i, \\ d(\boldsymbol{\psi}_i, \boldsymbol{\psi}_j) &= \boldsymbol{\psi}_i^T (\boldsymbol{\Psi}_{\text{old}}^T \boldsymbol{\Psi}_{\text{old}})^{-1} \boldsymbol{\psi}_j. \end{aligned} \tag{4.22}$$

Consequently, when the D-optimal design is constructed by maximizing $|\boldsymbol{M}_{\boldsymbol{\Psi}}|^{1/P}$, the couple $(\boldsymbol{\psi}_i, \boldsymbol{\psi}_j)$ which correspond to the largest value of the Fedorov's Δ-function are exchanged to provide the maximum improvement in the optimality criterion. The algorithm stops when the largest value of the Δ-function is smaller than a user specified threshold, e.g., 10^{-6} [84]. Convergence proof of the FEA can be found in [84].

At each iteration of the original FEA only one exchange is performed. A modified version of FEA with multipe exchanges in each iteration was proposed in [48]. Other popular exchange algorithms include DETMAX algorithm [162], Wynn's algorithm [256], k-exchange algorithm [120], kl-exchange algorithm [10], and coordinate exchange algorithm of [158] which does not require a candidate set. Performance comparison between different exchange algorithms can be found in [48, 120, 187, 172]. A discussion on details of these methods is beyond the scope of the present work and we refer the interested reader to the references provided for further details.

For either high-dimensional problems, or the cases in which high order PCEs are required, the number of coefficients to be estimated, P, may be very large. Consequently, the candidate measurement matrix should have a large number of rows to properly cover the design space Ω. This results in a large number of exchanges to be performed and eventually high computational costs. Moreover, for criteria other than

D-optimality, there is no update formula similar to (4.21) for the simultaneous inclusion and exclusion of the design points that can be used to expedite the exchange procedure.

In order to relax the computational complexity of constructing the alphabetic optimal designs, we employ a simple sequential (greedy) technique in this work. The sequential algorithm starts with an empty design and then one row from the candidate design corresponding to the largest improvement in the optimality criterion is added to design. This procedure continues until a design with N rows is constructed. As there is no exchange involved, this algorithm is considerably faster than FEA. Moreover, since only rows are being added to the design in sequential construction of the ODE, one may employ the simple update formula for the trace of inverse of matrix in D and expedite the construction of A-optimal designs obtained via the sequential algorithm.

Algorithm 5 A sequential algorithm to construct alphabetic optimal designs. Adopted from [68, 219, 230].

• **Input**: Number of samples in the design N, number of samples in the candidate design N_c, orthogonality measure $f(\boldsymbol{\xi})$, optimality measure ϕ.
• **Output**: $N \times P$ optimal measurement matrix $\boldsymbol{\Psi}$ and the corresponding optimal sample set $\{\boldsymbol{\Xi}^{(i)}\}_{i=1}^{N}$ based on the optimality measure ϕ.
Generate N_c realizations of $\boldsymbol{\xi}$, $\{\boldsymbol{\Xi}^{(i)}\}_{i=1}^{N_c}$, and the corresponding $N_c \times P$ candidate measurement matrix $\boldsymbol{\Psi}_c$.
Randomly select a row from $\boldsymbol{\Psi}_c$ to construct an initial $1 \times P$ design $\boldsymbol{\Psi}$.
for $i = 1 : N - 1$ **do**
    for $j = 1 : N_c - i$ **do**
        Compute the new optimality measure after adding a row from the remaining rows in $\boldsymbol{\Psi}_c$ to $\boldsymbol{\Psi}$, i.e., $\phi(j) = \phi\left(\boldsymbol{\Psi} \cup \boldsymbol{\Psi}_c(j,:)\right)$.
    end for
    Add a row from $\boldsymbol{\Psi}_c$ to $\boldsymbol{\Psi}$ that corresponds to the smallest optimality measure $\phi(j)$.
end for

A similar greedy scheme based on the sequential addition of design rows was first employed in [68] to construct D-optimal designs. This technique has also been employed in [230] to construct D- and A-optimal Bayesian designs. It was suggested that the sequentially generated optimal designs may be improved further by FEA, but a noticeable improvement was not observed by authors in the context of network performance inference [230]. A similar sequential algorithm has been used in [219] to construct a set of quasi-optimal samples by maximizing an optimality measure as a function of $|\boldsymbol{M}_{\boldsymbol{\Psi}}|$, i.e., $\mathcal{S}(\boldsymbol{\Psi}) := \left(\sqrt{|\boldsymbol{M}_{\boldsymbol{\Psi}}|} / \prod_{j=1}^{P} \|\boldsymbol{\Psi}^j\|_2\right)^{1/P}$, where $\boldsymbol{\Psi}^j$ is jth column of $\boldsymbol{\Psi}$, for the weighted LSR problem. Very recently, an accurate estimation of the PCE-based Sobol' indices via the LSR technique has been reported in [30] where the D-optimal sample set is generated using a similar greedy approach we employ

in this work. Here we employ this sequential scheme to construct A- and E-optimal designs as well as D-optimal designs.

Algorithm 5 presents the main steps in the construction of optimal designs via the sequential technique used in this study. The following points are worth highlighting regarding this algorithm:

- As we are only adding new rows to the design in Algorithm 5, one may expedite the computation of D- and A-optimality measures by employing the update formulas given in (C.1) and (D.1), respectively. There is no update formula for the E-optimal designs, hence, their construction is considerably more expensive in comparison to the other two optimality criteria.

- In computer experiments, including regression-based PCEs, unlike the real physical experiments, replication of the samples is not allowed as there is no measurement noise in computer experiments and their outputs are deterministic. Hence, it is crucial to search among the remaining rows of the candidate measurement matrix $\mathbf{\Psi}_c$ to avoid having replicated rows in $\mathbf{\Psi}$.

- For the case of D-optimal designs, when the number of rows in $\mathbf{\Psi}$ is smaller than P, the determinant of the corresponding information matrix is zero. To overcome this issue, a column-truncated version of the candidate measurement matrix is here considered, where rows and columns are added to the design simultaneously until a $P \times P$ design is constructed. After that, only rows are added to the design. We refer the readers to [219, Section 3.3.1] for more details on this matter.

- In general, constructing the exact alphabetic optimal designs is known to be an NP-hard optimization problem [151], thus finding a global optimal design becomes infeasible as the dimension of the problem increases. Hence, an optimal design obtained by greedy search algorithms, such as Algorithm 5, is in fact quasi-optimal rather than being globally optimal. Moreover, the alphabetic optimal criteria, ϕ, are not convex and typically exhibit multiple local minima/maxima. In order to increase the chance of finding the global optima, it is advised to run them multiple times with different (random) initial designs.

- We note that according to the discussion in Remark 8, the resulting A-optimal design for the or-

thogonal PCEs is also I-optimal.

In addition to the greedy algorithms discussed here, other techniques such as simulated annealing (SA) [106, 157], genetic algorithms (GA) [163, 27], semi-definite programming (SDP) [199, 8] and pattern search [112, 109] have been employed in the literature to construct alphabetic optimal designs. Moreover, software packages such as JMP [101], POBE [154], AlgDesign [253], PopED [179, 87] and Gosset [110] have been developed to construct ODE using the algorithms discussed here. We note that the majority of the software packages for the construction of ODE are developed for scenarios where low-order (mostly quadratic) polynomials are employed, hence, limiting the applicability of such packages for high-order PCEs.

4.4.6 Space-filling designs

Space-filling designs form a class of sampling techniques that have received extensive attention in the context of computer simulation experiments [135, 209, 188, 125]. As mentioned in Section 4.4.5.1, computer experiments are deterministic and their output QoI is free of the measurement noise, and it is usually assumed that the true response trend between the inputs and outputs of a computer model is not known a priori. Designs capable of covering all of the portions of the design space $\mathbf{\Omega}$, hence the name space-filling, have been proposed to reduce the error in computer experiments. In the following, we will review two important space-filling designs, namely quasi-Monte Carlo (QMC) and Latin hypercube (LH) sampling techniques.

4.4.6.1 Quasi-Monte Carlo sampling

To alleviate the slow convergence rate of standard MC technique, especially for high-dimensional integrations in finance applications [252], researchers have proposed a deterministic version of the MC sampling method based on the *low-discrepancy* sequences [54] known as *quasi-Monte Carlo* (QMC) technique [32, 173, 159, 24]. It has been shown that for d-dimensional uniform random variables, a convergence rate of approximately $\mathcal{O}(N^{-1}(\log N)^d)$, which is asymptotically better than that of MC, can be expected from

the QMC technique [32].

Unlike *pseudo-random* sequences used in standard MC, low-discrepancy sequences, a.k.a. *quasi-random*, such as Hammersley set, Niederreiter, Halton, Faure, and Sobol sequences [156], are constructed such that the maximal degree of uniformity is provided for the sample points drawn from a uniform distribution. A quantitative assessment of the degree of uniformity of a set of sample points is achieved by calculating its *discrepancy*. Among the many definitions proposed for the discrepancy of a sequence, star discrepancy D_N^*, defined in Section 4.3.2, is one of the most widely used ones due to its simplicity.

In [159], a stable and accurate solution to the LSR problem in (4.4) is achieved with low-discrepancy sample points and a number of samples N which is proportional to the square of P (up to a logarithmic factor). To the best of our knowledge, the application of QMC sampling technique for regression-based PCEs has not been extended to problems with non-uniform random inputs. For a detailed description of QMC sampling approach for multivariate Legendre polynomials we refer the interested reader to [159].

Another class of space-filling designs, known as *uniform* design, which facilitate the same idea of constructing samples with maximum degree of uniformity was proposed in [75]. Uniform designs are constructed based on alternative and computationally less expensive measures of the uniformity rather than the star discrepancy D_N^* used in QMC scheme [74]. Several methods such as good lattice method [78], optimization searching method [255], collapsing method [76], combinatorial construction method [269], and Latin square method [77] have been proposed to construct uniform designs.

4.4.6.2 Latin hypercube sampling

Latin hypercube (LH) sampling, first introduced in [145], aims at selecting the sample set $\{\boldsymbol{\Xi}^{(i)}\}_{i=1}^{N}$ such that these points cover all portions of the design space. LH sampling attempts to reduce the variance of the estimator by selecting N sample points from N equiprobable portions of the random space [97], which is known as *stratification* of the probability distribution. In [232, 181], it is shown that the convergence rate of the LH sampling is never worse than that of standard MC when $N \to \infty$. However, under some conditions discussed in [145], e.g., when certain monotonicity conditions do not hold, LH may not be advantageous over the standard MC in reducing the variance of the estimator. Algorithm 6 summarizes the main steps to

generate $\{\boldsymbol{\Xi}^{(i)}\}_{i=1}^{N}$ via LH sampling technique.

Algorithm 6 Summary of the Latin hypercube sampling approach.

1: Divide the PDF of each random variable $\{\xi_k\}_{k=1}^{d}$ into N equiprobable intervals.
2: Compute the sampled cumulative probability in the ith interval as $\text{Prob}_i = \frac{i-1}{N} + \frac{r_u}{N}$, where r_u is a uniform random number drawn independently from $[0, 1]$.
3: Compute the ith realization of random variable ξ_k, $\Xi_k^{(i)}$, via the inverse of its cumulative distribution function (CDF), i.e., $\Xi_k^{(i)} = F_{\xi_k}^{-1}(\text{Prob}_i)$.
4: Pair the N realizations of each random variable ξ_k with the N realizations of other variables.

Several modified versions of LH sampling such as orthogonal [265, 31, 266] and column-wise orthogonal LH designs [233], sliced and optimal sliced LH designs [190], and orthogonal-maximin LH designs [127] have been proposed to improve the efficiency and accuracy of the LH sampling approach. Algorithm 6 is used when the random inputs $\{\Xi_k\}_{k=1}^{d}$ are independent from each other. For the correlated random inputs, some extended/modified versions of LH sampling have also been proposed in the literature [114, 115, 201].

In addition to the space-filling designs we reviewed here, many other have been introduced over the past few decades. Among them we mention maximin designs [119], a.k.a. sphere packing designs, in which the minimum distance between pairs of designs points is maximized, maximum entropy designs [217], which takes the entropy as a measure of the uniformity of the samples, Gaussian process integrated mean squared-error (IMSE) designs minimizing the variance of prediction [117, 198], and recently maximum projection designs [126].

Remark 10. *Although alphabetic optimal designs are developed for real world physical experiments, but these designs have also been successfully employed for computer experiments. As an example, we mention the work in [122], which compares D- and I-optimal designs with space-filling designs such as Latin hypercube, uniform, sphere packing, and maximum entropy, in terms of reducing the prediction variance for high-order polynomials. It was shown that D- and I-optimal designs result in prediction variances smaller than those obtained by space-filling designs. A similar comparison was also made in [121].*

Next, we explore the performance of some selected sampling techniques when applied to different stochastic problems.

4.5 Numerical Examples

In this section, we study the empirical performance of six different sampling methods by considering three numerical examples. In the first example, the recovery of randomly generated functions is investigated. Next, a nonlinear Duffing oscillator under free vibration which requires high order PC expansions is considered. And finally, we examine the application of a subset of the reviewed sampling techniques to the stochastic remaining useful life (RUL) prediction of Lithium-ion batteries (LIBs).

We consider standard MC (Section 4.4.1), coherence-optimal (Section 4.4.2) and Latin hypercube (Section 4.4.6.2) techniques denoted by *Standard*, *coh-opt* and *LH* sampling, respectively. In addition, we introduce hybrid sampling methods, denoted by *alphabetic-coherence-optimal*, that combine coherence-optimal technique with alphabetic optimality criteria. To construct these sample sets, the N_c realizations of $\boldsymbol{\xi}$ in the third line of Algorithm 5 are generated via the coherence-optimal technique of Section 4.4.2. Consequently, the corresponding $N_c \times P$ candidate measurement matrix $\boldsymbol{\Psi}_c$ to be used in the sequential construction of D-, A-, and E-optimal designs will be already optimal in a sense that it yields to a small coherence parameter μ, hence the names *D-coh-opt*, *A-coh-opt*, and *E-coh-opt*, respectively. We note that the larger N_c in Algorithm 5, the higher the quality of the sequentially constructed design and the associated computational cost.

To verify the accuracy of the regression-based PC model constructed with N independent samples, we compute the relative error of the resulting PC expansion in predicting independent realizations of the QoI as

$$\text{relative error} = \frac{\|\boldsymbol{u}_v - \boldsymbol{\Psi}_v \boldsymbol{\alpha}\|_2}{\|\boldsymbol{u}_v\|_2}, \tag{4.23}$$

where $\boldsymbol{u}_v$ is the vector of N_v additional realizations of QoI (not used in computing $\boldsymbol{\alpha}$) and $\boldsymbol{\Psi}_v$ is the measurement matrix corresponding to $\boldsymbol{u}_v$.

Remark 11. *Because of the random nature of the discussed sampling techniques, the results presented in the following are obtained by averaging 60 independent replications using N samples to compute $\boldsymbol{\alpha}$ and N_v realizations to evaluate the relative error in (4.23).*

4.5.1 Manufactured functions

In the first example, we consider the reconstruction of manufactured PC expansions, where the coefficients $\boldsymbol{\alpha}$ are prescribed to generate $\boldsymbol{u}$. Specifically, we consider the model

$$u(\boldsymbol{\Xi}^{(i)}) = \sum_{j=1}^{P} \alpha_j \psi_j(\boldsymbol{\Xi}^{(i)}), \qquad i = 1, \cdots, N, \tag{4.24}$$

where each α_j is generated independently from a standard normal distribution. Each realization $u(\boldsymbol{\Xi}^{(i)})$ is then generated by evaluating the right-hand-side of (4.24) at a set of N samples of $\boldsymbol{\xi}$. Additionally, we assume that there is independent, normally distributed additive noise in the evaluation of $u(\boldsymbol{\Xi}^{(i)})$ with a standard deviation of $0.03 \cdot |u(\boldsymbol{\Xi}^{(i)})|$. The PC expansion is then recovered successfully if the least squares solution yields a relative error ≤ 0.02.

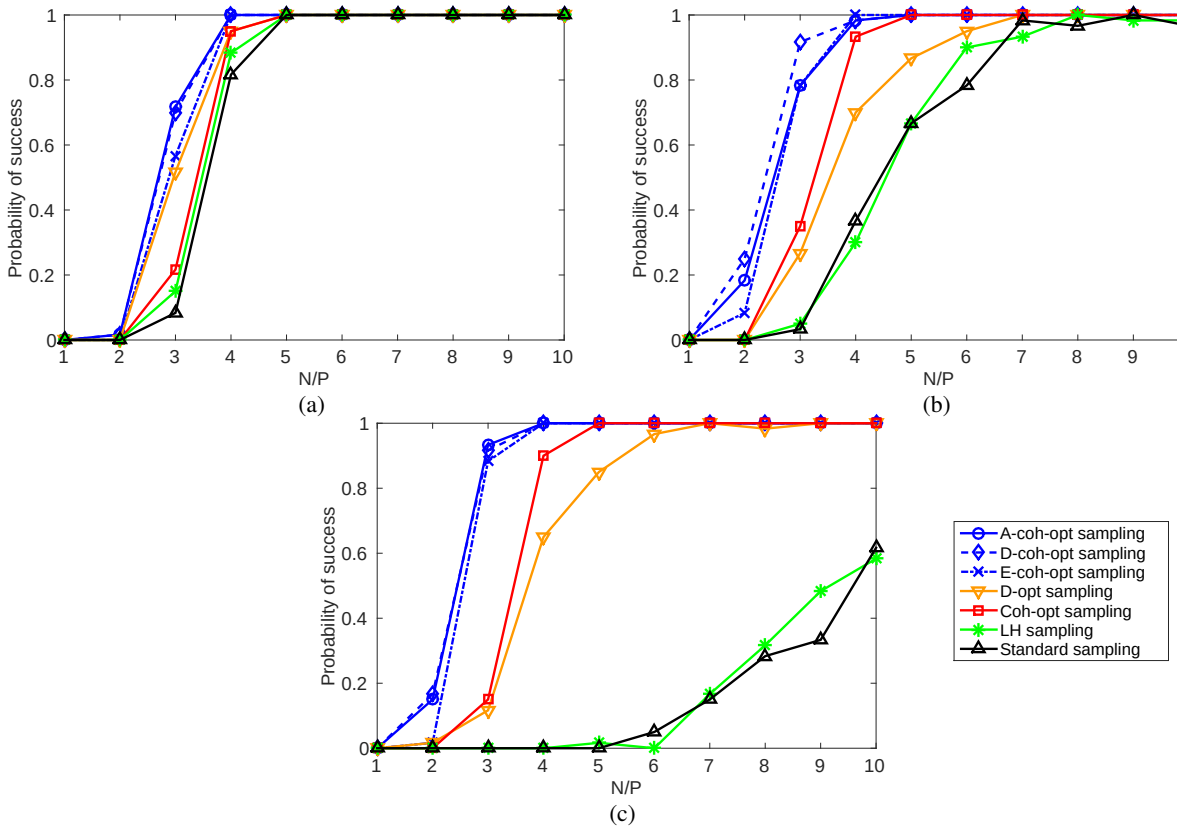


Figure 4.1: Legendre recovery probability as a function of sample size N. (a) $(p, d) = (2, 15)$; (b) $(p, d) = (4, 4)$; (c) $(p, d) = (15, 2)$.

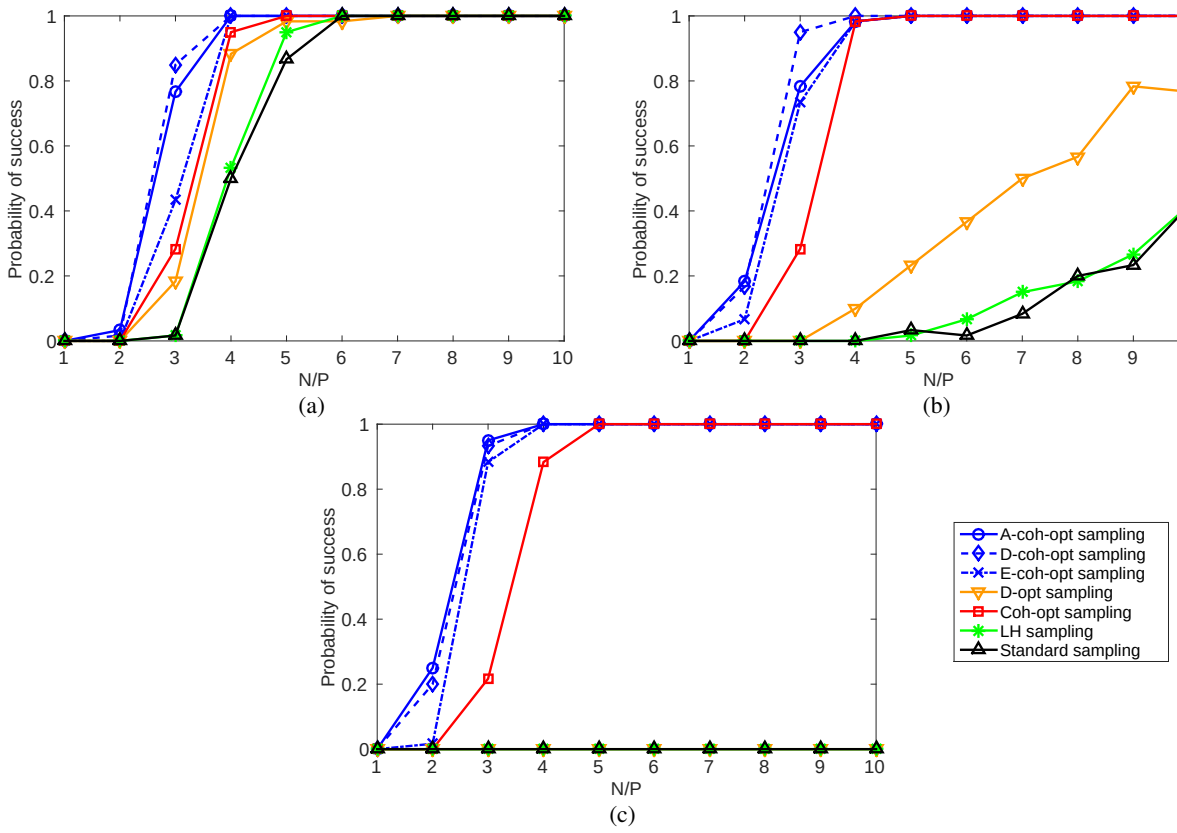


Figure 4.2: Hermite recovery probability as a function of sample size N. (a) $(p, d) = (2, 15)$; (b) $(p, d) = (4, 4)$; (c) $(p, d) = (15, 2)$.

Figs. 4.1 and 4.2 represent the probability of a successful recovery of the manufactured solution as a function of the number of samples N to the number of unknowns P for the case of Legendre and Hermite polynomials, receptively. Three different scenarios, i.e., a high-dimensional case with $(p, d) = (2, 15)$, a moderate case of $(p, d) = (4, 4)$, and a high-order case with $(p, d) = (15, 2)$ are considered in these figures. In this example, we consider $N_c = 4N$ for each value of the oversampling ratio N/P. The following observations are worth highlighting regarding Figs. 4.1 and 4.2:

- For the high-order case $(p, d) = (15, 2)$, both standard and LH sampling techniques demonstrate a poor performance in recovering the solution such that for the case of Hermite polynomials, they both fail to recover the solution even for large sample sizes, e.g., $N = 10P$. These two methods perform significantly better for the high-dimensional, low-order case $(p, d) = (2, 15)$, but are yet

worse than the other sampling techniques for the moderate case $(p, d) = (4, 4)$.

- Coherence-optimal approach results in a more accurate solution recovery in comparison to the standard and LH sampling methods, especially, for the high-order case $(p, d) = (15, 2)$.

- In all cases, alphabetic-coherence-optimal techniques outperform the other sampling methods. They are significantly better than the standard and LH techniques for the high-order case $(p, d) = (15, 2)$. Their advantage becomes more significant when the oversampling ratio N/P is small.

- Among the alphabetic-coherence-optimal techniques, *D-coh-opt* is slightly better than *A-coh-opt*, while *E-coh-opt* is the worst of three. To justify this, we note that E-optimal designs only consider the largest error of the estimator, while the geometric and arithmetic mean of the errors are considered in D- and A-optimal designs, respectively. Given that, it is generally accepted that E criterion uses less information in comparison to D and A criteria and provides less accurate results [18].

- The main difference between the alphabetic-coherence-optimal technique and the alphabetic optimal design is that in the former approach the candidate ensemble $\mathbf{\Psi}_c$ is generated based on the coherence-optimal technique while in the latter $\mathbf{\Psi}_c$ is constructed via random sampling from the orthogonality measure $f(\mathbf{\Xi})$. To highlight the importance of constructing $\mathbf{\Psi}_c$ via coherence-optimal approach, we also report the results of the classical D-optimal design, denoted by *D-opt* in Figs. 4.1 and 4.2, in which the candidate measurement matrix of Algorithm 5 is constructed by drawing samples from $f(\mathbf{\Xi})$ randomly. Although *D-opt* sampling performs fairly good for the the high-dimensional case in comparison to the alphabetic-coherence-optimal techniques, its accuracy is drastically reduced when the total order of the expansion, p, is increased.

4.5.2 Nonlinear Duffing oscillator

The second problem of interest in this study is to quantify the uncertainty in the displacement solution $u(\boldsymbol{\xi}, t)$ of a nonlinear single degree-of-freedom (DOF) Duffing oscillator [147] under free vibration

described by

$$\begin{aligned} &\ddot{u}(\boldsymbol{\xi},t) + 2\omega_1\omega_2\dot{u}(\boldsymbol{\xi},t) + \omega_1^2(u(\boldsymbol{\xi},t) + \omega_3 u^3(\boldsymbol{\xi},t)) = 0, \\ &u(\boldsymbol{\xi},0) = 1, \qquad \dot{u}(\boldsymbol{\xi},0) = 0, \end{aligned} \tag{4.25}$$

where the uncertain parameters $\{\omega_i\}_{i=1}^3$, describing the motion of the oscillator, are given by

$$\begin{aligned} \omega_1 &= 2\pi\ (1 + 0.2\ \xi_1), \\ \omega_2 &= 0.05\ (1 + 0.05\ \xi_2), \\ \omega_3 &= -0.5\ (1 + 0.5\ \xi_3), \end{aligned} \tag{4.26}$$

with $\{\xi_i\}_{i=1}^3$ being i.i.d. uniform random variables $U(-1,1)$.

Although the stochastic dimensionality of this problem is not high, i.e., $d = 3$, due to the presence of nonlinearities, high order PC expansions are required in order to preserve the accuracy of approximation for large instances of t. This high-order approximation requirement is highlighted in Fig. 4.3 where we report the mean and standard deviation of the relative error of the standard MC sampling approach in estimating the displacement at $t = 4$ seconds as a function of the total order of expansion p. In this experiment, we assumed $N = 20P$ to make sure that we have enough samples for the cases with large p values and considered an additional $N_v = 10000$ validation samples in (4.23) to compute the relative error.

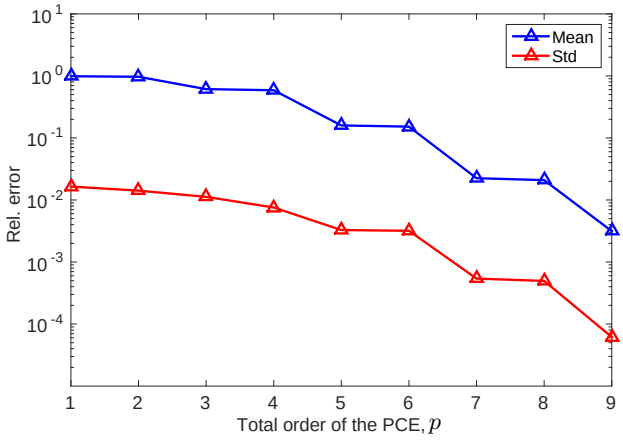


Figure 4.3: Mean and standard deviation of the relative error of standard sampling with $N = 20P$ samples in estimating the displacement $u(\boldsymbol{\Xi}, 4)$ versus the total order of PCE.

We considered a high order PCE, i.e., $p = 9$, for this problem to examine the performance of sampling

techniques. The pair $(p, d) = (9, 3)$ corresponds to $P = 220$ unknown PC coefficients to be approximated. Since P is large, and we need to have $N \geq P$ for an stable LSR solution, the issue of choosing the best sampling approach for this problem becomes critical.

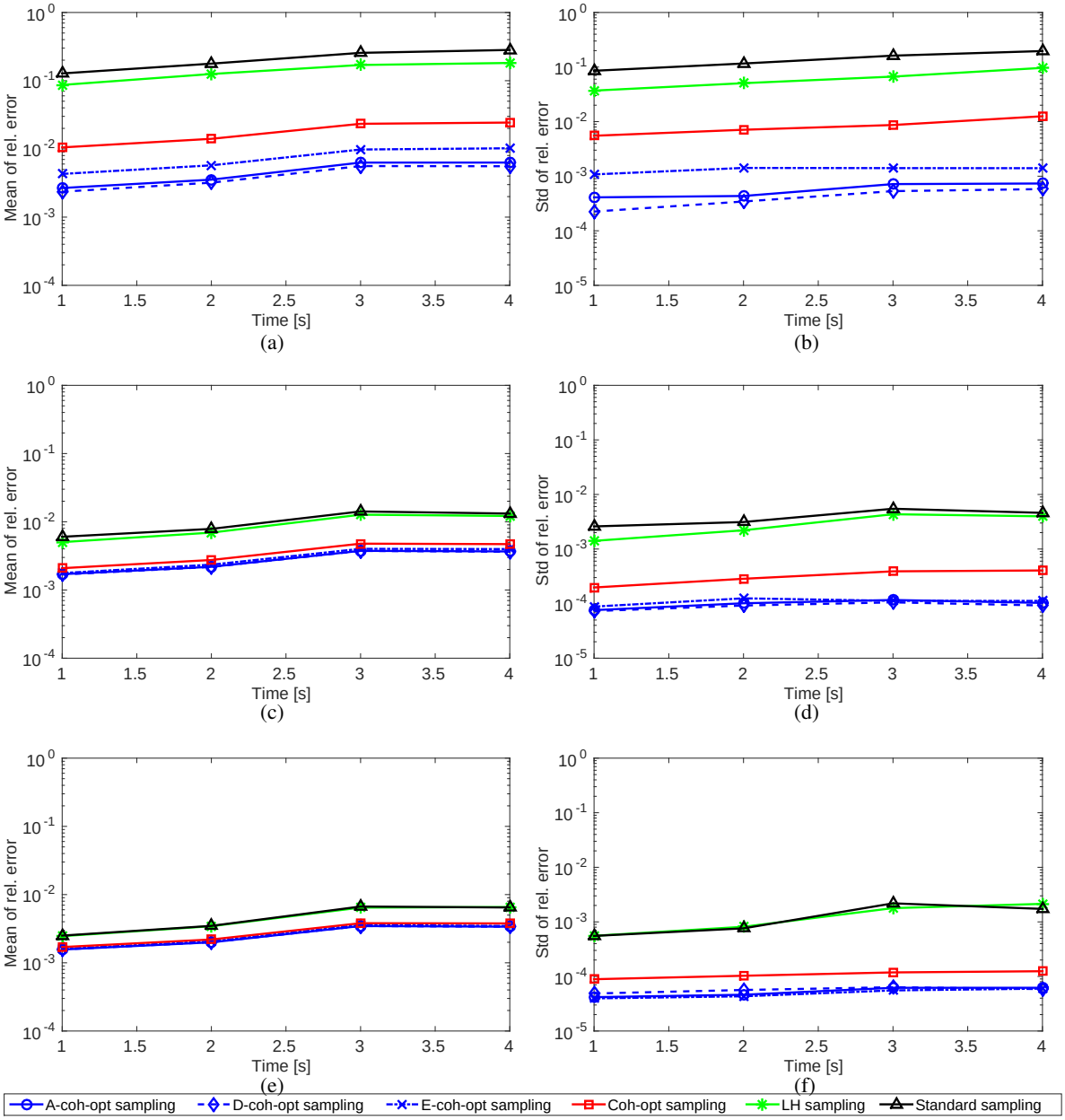


Figure 4.4: Mean and standard deviation of the relative error in estimating the displacement $u(\Xi, t)$ with a 9th order PCE and: (a)-(b): $N = 242$, (c)-(d): $N = 440$, (e)-(f): $N = 660$.

Fig. 4.4 demonstrates the mean and standard deviation of the relative error in predicting $N_v = 10000$

independent realizations of the displacement obtained by different sampling methods. Three oversampling ratios of 1.1, two and three which correspond to $N = 242, 440$ and 660 samples, respectively, have been considered. Following [168], we assume $N_c = \lfloor 1.5P\log(P)\rfloor$, where P is the number of unknown PC coefficients and is given in (1.5).

Again, the long-time integration issue we discussed earlier is observed here as well; that is independent of the sampling technique, the relative error is slightly increased over time when p is kept constant. One possible approach to address this issue is to employ Multi-Element generalized Polynomial Chaos (ME-gPC) in which the total order of expansion is adjusted over the time. Discussion on ME-gPC is beyond the scope of this paper and we refer the interested reader to [248] on this subject.

As it is shown in Fig. 4.4, the alphabetic-coherence-optimal techniques result in smaller values for the relative error in comparison to the other methods, with *D-coh-opt* and *A-coh-opt* to perform slightly better than *E-coh-opt*. This superiority is signified in Figs. 4.4a-b where an oversampling ratio of 1.1 is used. These plots suggest that for the case of high-order PC expansions, the alphabetic-coherence-optimal techniques are the most suitable sampling methods, in particular, when the number of samples is slightly larger than the number of unknown coefficients P. For the applications where evaluation of the QoI is computationally expensive, we prefer to approximate the PCE with a minimum number of solution realizations, usually with N slightly larger than P.

The standard MC and LH techniques provide the least accurate results, specially, for the case of $N = 242$ where the oversampling ratio is low. In this case, these two methods fail to accurately approximate the solution as the error is not only large, but it is also highly variable as suggested by Fig. 4.4b. As the number of samples is increased, MC and LH methods perform considerably better such that with $N = 440$ samples, both techniques result in acceptable relative errors.

Similar to the previous example, the relative error obtained by the coherence-optimal approach falls somewhere between the error reported by MC and LH methods and the alphabetic-coherence-optimal techniques. By comparing Figs. 4.4a, 4.4c, and 4.4e, it can be seen that the coherence-optimal results converge to the alphabetic-coherence-optimal results at a higher rate in comparison to MC and LH methods as N is increased.

4.5.3 Stochastic prediction of remaining useful life of batteries

The last problem of interest in this study is the model-based stochastic estimation of the remaining useful life (RUL) of Lithium-ion batteries (LIBs). RUL of LIBs is defined as the amount of time before the battery health falls below a defined threshold, e.g., reaching the cut-off potential at the end of discharge [202], and is studied under the context of prognostics and health management (PHM) [239]. For the sake of completeness, a brief introduction to model-based prognostics of LIBs with an emphasis on RUL predictions is presented in the following.

4.5.3.1 Model-based prognostics

In general, prognosis approaches may be categorized as data-driven and model-based [14, 206]. A mathematical representation of the system based on the physics of the problem is used in model-based methods while data-driven techniques use system monitoring data instead [206]. Both techniques have been successfully employed to predict the RUL of batteries [220, 202, 143, 142, 51, 204]. Here, we focus on model-based RUL estimation for LIBs which is suitable for online health monitoring.

Let us assume that the system model is given by the following set of equations,

$$\dot{\boldsymbol{z}}(t) = f(t, \boldsymbol{z}(t), \boldsymbol{\theta}(t), \boldsymbol{\nu}(t), \boldsymbol{v}_p(t)), \tag{4.27}$$

$$\boldsymbol{y}(t) = h(t, \boldsymbol{z}(t), \boldsymbol{\theta}(t), \boldsymbol{\nu}(t), \boldsymbol{v}_m(t)), \tag{4.28}$$

where $\boldsymbol{z}(t)$ is the state vector, f is the state equation, $\boldsymbol{\theta}(t)$ is the model parameter vector, $\boldsymbol{\nu}(t)$ is the input vector, $\boldsymbol{v}_p(t)$ is the process noise vector, $\boldsymbol{y}(t)$ is the output vector, h is the output equation, and $\boldsymbol{v}_m(t)$ is the measurement noise vector.

As illustrated in [202, Fig. 1], the computational framework for prognostics and RUL predictions includes three main steps: estimation of the state $\boldsymbol{z}$ at time t_p, prediction of the future states, and finally RUL computation. Kalman or particle filtering approaches may be employed within a Bayesian framework to estimate the state vector $\boldsymbol{z}$ using the output data measured up to time t_p. Both Eqs. (4.27) and (4.28) should be used in the estimation step. Following [202], in this study we assume that the estimated state vector at time t_p is given and we focus on the prediction and RUL computation steps.

Having at hand the estimated $\boldsymbol{z}$ up to time t_p, predicting the future states of the system is the next step to compute the RUL. In the prediction step, only the state space model in (4.27) is used since no measured data is available for $t > t_p$. Here, the goal is to predict the occurrence of an event E defined with respect to the states, parameters, and inputs of the system [50]. Using a set of m constraints $\{\kappa_i\}_{i=1}^{m}$, one may define an acceptable region of performance for the system. The constraint $\kappa_i : \mathbb{R} \to \mathbb{B}$ maps a given point in the joint state-parameter space given the current inputs, i.e., $(\boldsymbol{z}(t), \boldsymbol{\theta}(t), \boldsymbol{\nu}(t))$, to a Boolean domain $\mathbb{B} := \{0, 1\}$, where $\kappa_i = 1$ if the constraint is satisfied, and zero otherwise [202]. One may combine these individual constraints into a single threshold function $T_{EOL} : \mathbb{R} \to \mathbb{B}$ as

$$T_{EOL}(\boldsymbol{z}(t), \boldsymbol{\theta}(t), \boldsymbol{\nu}(t)) = \begin{cases} 1, & 0 \in \{\kappa_i(\boldsymbol{z}(t), \boldsymbol{\theta}(t), \boldsymbol{\nu}(t))\}_{i=1}^{m}, \\ 0, & \text{otherwise}, \end{cases} \tag{4.29}$$

where the subscript EOL denotes the end of life. In (4.29), violation of any constraints results in $T_{EOL} = 1$. Then, one may define the end of life at time t_p, $E(t_p)$, as the earliest time at which $T_{EOL} = 1$, i.e.,

$$E(t_p) := \inf\{t \in \mathbb{R} : t \geq t_p \wedge T_{EOL} = 1\}. \tag{4.30}$$

Given $E(t_p)$, the RUL at time t_p, $R(t_p)$, is then given by

$$R(t_p) := E(t_p) - t_p. \tag{4.31}$$

For the case of LIBs, the event E we are interested in predicting is the end of discharge. T_{EOL} is then specified as $V < V_{\text{cutoff}}$, where V_{cutoff} is a user specified cut-off voltage and V is the terminal voltage of the battery. Hence, the only constraint in predicting the RUL of LIBs is $\kappa_1 : V > V_{\text{cutoff}}$.

4.5.3.2 An electrical equivalent circuit model for LIBs

In order to facilitate the application of the model-based RUL estimation for the online health monitoring of LIBs, the mathematical model described by Eqs. (4.27) and (4.28) is desired to be computationally inexpensive. Hence, we consider an empirical model of the LIB, which captures its major dynamics, except the temperature effects, based on the electrical equivalent circuit shown in Fig. 4.5. This model has been employed in [202, 203, 205] to simulate a battery used to power an unmanned aerial vehicle. Charge of the

LIB q_b is held by the capacitance C_b. The nonlinear C_b captures the open-circuit potential and concentration over-potential. The non-linear voltage drop due to surface over-potential is captured by $R_{sp} - C_{sp}$, R_s models the ohmic drop, and R_p represents the parasitic resistance that accounts for self-discharge [203].

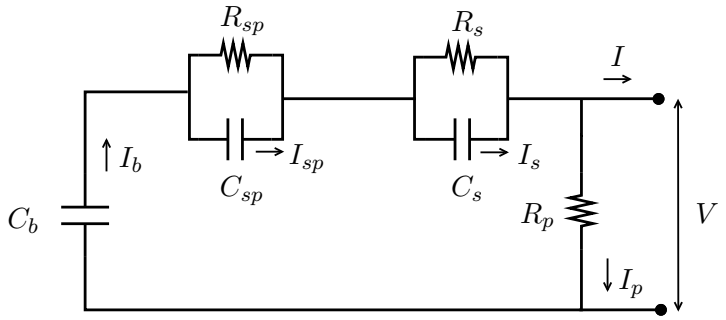


Figure 4.5: LIB equivalent circuit model [205].

Denoting the LIB maximum capacity and charge by C_{max} and q_{max}, respectively, the LIB state of charge, SoC, is given by

$$SoC = 1 - \frac{q_{max} - q_b}{C_{max}}. \tag{4.32}$$

The concentration polarization resistance R_{sp} is a non-linear function of SoC,

$$R_{sp} = R_{sp0} + R_{sp1} \exp[R_{sp2}(1 - SoC)], \tag{4.33}$$

where R_{sp0}, R_{sp1}, and R_{sp2} are empirical parameters [203]. C_b is also given by

$$C_b = C_{b0} + C_{b1}SoC + C_{b2}SoC^2 + C_{b3}SoC^3, \tag{4.34}$$

with C_{bi}, $i = 0, 1, 2, 3$, being constants determined empirically. The model parameters of the considered cell are shown in Table 4.1.

Voltage drops across the individual circuit elements and their associated currents and charges are governed by the set of equations presented in Table 4.2 where I denotes the battery discharge current. The terminal voltage of the cell V is then given by $V = V_b - V_{sp} - V_s$. We assume that in this example $V_{\text{cutoff}} = 16$V, hence, $\kappa_1 : V > 16$.

Different uncertainty sources such as uncertainty in the future inputs, state, model and process noise have been identified for the stochastic RUL estimations [203, 50, 206, 51, 202, 200, 205]. We refer the

Table 4.1: LIB model parameters [205].

Parameter	Value
R_{sp0}	0.0272 Ω
R_{sp1}	1.087×10^{-16} Ω
R_{sp2}	34.64
R_s	0.0067 Ω
R_p	10000 Ω
C_{b0}	19.8 F
C_{b1}	1745 F
C_{b2}	-1.5 F
C_{b3}	-200.2 F
C_s	115.28 F
C_{sp}	316.69 F
q_{max}	31100 C
C_{max}	30807 C

readers to the provided references for more details on the sources of uncertainty for this problem. In this example, we assume that the battery is discharged at a constant current (in Amps) represented by a random variable following a Beta distribution with shape parameters $\alpha = 21.2$ and $\beta = 31.8$. As mentioned earlier, we assume that the state estimates are already obtained via a Bayesian framework. For this example, we consider an estimation of the state variables, q_b, q_{sp}, and q_s, given in [202, Figs. 5-7], respectively, with a coefficient of variation (COV) of 0.1. Whenever the estimated state is equal to zero, we assume its standard deviation is equal to 0.1. In addition, three process noise terms with zero mean normal distributions and variances of 0.1, 10^{-4}, and 10^{-6} are considered in the state equations describing q_b, q_{sp}, and q_s, respectively. Assuming that there is no model uncertainty, the total stochastic dimensionality of this problem is then $d = 7$. We note that the corresponding PC basis function for this problem consists of Jacobi and Hermite polynomials which correspond to Beta and normal random variables, respectively.

Fig. 4.6 shows the relative error of the sampling techniques in estimating the RUL of LIB at $t = 0, 200, 400$, and 600 seconds. Similar to the previous example, 10000 independent samples are used to

Table 4.2: Equations describing the electrical equivalent circuit model for LIBs [205].

	Governing equation
Voltages	$V_b = q_b/C_b, \quad V_{sp} = q_{sp}/C_{sp}, \quad V_s = q_s/C_s, \quad V_p = V_b - V_{sp} - V_s$
Currents	$I_p = V_p/R_p, \quad I_b = I_p + I, \quad I_{sp} = I_b - V_{sp}/R_{sp}, \quad I_s = I_b - V_s/R_s$
Charges	$\dot{q_b} = -I_b, \quad \dot{q}_{sp} = I_{sp}, \quad \dot{q_s} = I_s$

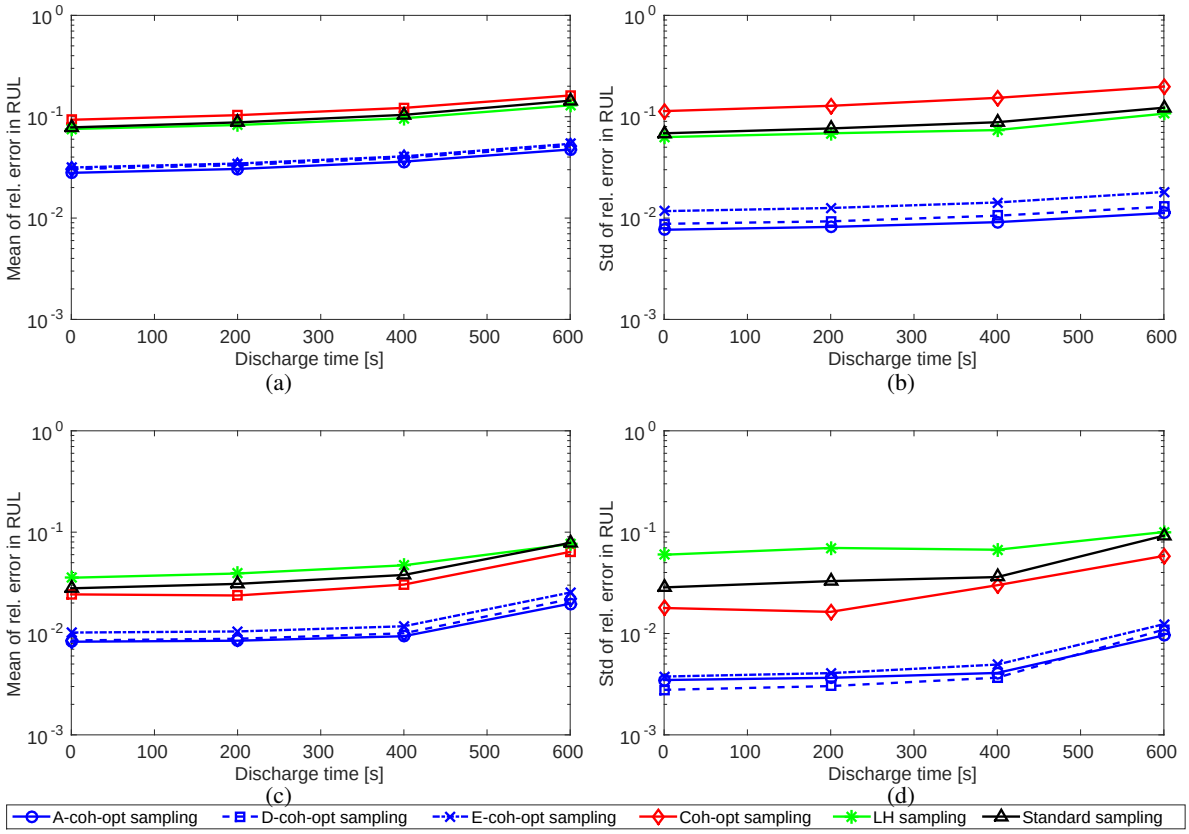


Figure 4.6: Mean and standard deviation of the relative error in estimating the RUL; (a)-(b): $p = 2$, $P = 36$, $N = 37$, (c)-(d): $p = 3$, $P = 120$, $N = 121$.

compute the relative error. Since model-based RUL predictions are mainly used in online applications where fast computations are critical, we limit the number of sample points in this example to be $N = P+1$, which corresponds to 36 and 121 samples for the second and third order PCEs, respectively, given $d = 7$. Similar to the previous example, we assumed $N_c = \lfloor 1.5P\log(P)\rfloor$. Fig. 4.6 demonstrates that similar to the high-order PC expansions, alphabetic-coherence-optimal techniques perform better that other methods for the high-dimensional problems as well when the oversampling ratio is low. Moreover, it can be seen that for the case of $d > p$ with low oversampling ratios, coherence-optimal method is not necessarily advantageous over the standard MC and LH samplings.

It is shown in Fig. 4.6c-d that the *D-coh-opt* approach results in the smallest errors. We next compare compare the PDF of the predicted RUL obtained by a third order PCE, obtained using the *D-coh-opt*

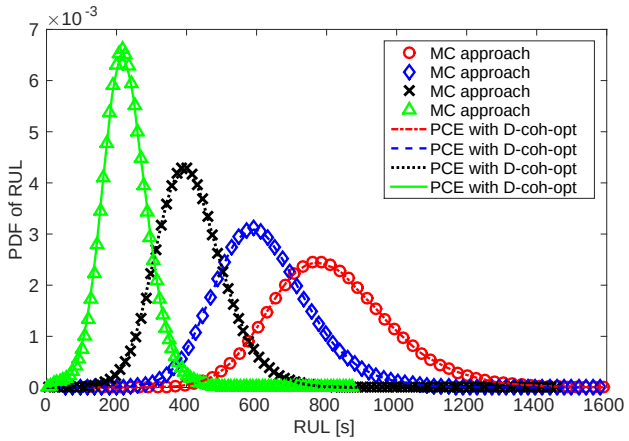


Figure 4.7: PDFs of the estimated RUL at $t = 0, 200, 400, 600$ seconds demonstrated in red, blue, black and green colors, respectively. *D-coh-opt* solution is obtained with a 3rd order PCE with $N = 121$ samples, while 20×10^6 used to obtain the MC solution.

approach, with a MC solution in Fig. 4.7. Although the results obtained by these two different approaches are overlapped, we note that 20×10^6 solution realizations are used in constructing the MC solution, while *D-coh-opt*-based PCE only requires 121 samples. This highlights the significant advantage of PCE approach equipped with optimal sampling methods over the traditional UQ techniques for the online RUL prediction of LIBs.

4.6 Summary and Conclusions

We provided a comprehensive review of the sampling techniques available in the literature for polynomial chaos expansions via least squares regression. We also proposed a hybrid sampling approach denoted by *alphabetic-coherence-optimal* technique in in which, given the same sample size, a combination of alphabetic optimal designs, such as *D*-, *A*-, and *E*-optimal, and the coherence-optimal technique results in more accurate approximations of the PC coefficients in comparison to other sampling methods.

In order to provide a road map for the practitioners seeking the best sampling method for their stochastic problem to be solved by regression-based PCE, we compared the empirical performance of the selected sampling techniques by their application to three stochastic problems with high-order and/or high-dimensional PCEs constructed with multidimensional Legendre, Hermite and Jacobi polynomials. Our

observations suggest the following considerations in selecting the most suitable sampling method for the regression-based PCE based on the total order of expansion p, stochastic dimension d and the oversampling ratio N/P:

- High-order PCE with $p > d$: We observed that for stochastic problems in which high-order expansions are required to capture an accurate solution, the alphabetic-coherence-optimal techniques deliver the most accurate approximations. We note that these sampling approaches are computationally more expensive in comparison to the other methods discussed in this study, but we emphasize that for the real world engineering applications where a singe realization of the output QoI may be very expensive, the extra computational cost for sampling is frequently acceptable and may be negligible. Among the alphabetic-coherence-optimal techniques, *D-coh-opt* results were slightly better than those obtained by *A-coh-opt*. In our experiments, we observed that the construction of the *D-coh-opt* sample sets is considerably faster than *A-coh-opt* samples, hence, we recommend *D-coh-opt* sampling technique for the high-order PC expansions.

- High-dimensional PCE with $d > p$: Although the alphabetic-coherence-optimal techniques still perform better than the other methods for the case of $d > p$, the standard MC and LH sampling methods may also deliver accurate approximations when low-order expansions are sufficient to approximate a high-dimensional QoI. In such cases, one may employ standard MC, LH, or other advanced space-filling sampling techniques discussed in this study as they may be less computationally expensive than the alphabetic-coherence-optimal approach.

- The oversampling ratio N/P: In the numerical examples provided in this study, it was observed when N is 3-4 times larger than the number of PC coefficients, one can expect that the standard MC and LH sampling methods exhibit a fairly acceptable performance in approximating the unknowns for moderate d and p. But as either of the stochastic dimensionality d or the expansion order p is increased, even with a fairly large oversampling ratio these methods may fail to deliver an accurate PC approximation. For the cases with small values of N/P, alphabetic-coherence-optimal techniques outperform other methods, specially when high-order PCEs are employed.

We also note that the sequential technique we used to construct the alphabetic-coherence-optimal samples may be substituted with more advanced exchange algorithms that may eventually improve the quality of the sample set further. Although, we note that for the high-dimensional and/or high-order polynomials that correspond to large P values, exchange algorithms may be computationally very expensive, hence, fast and efficient construction of the sample set based on the alphabetic optimality criteria will remain an open question in the context of non-convex optimization when the number of unknowns P is large.

Appendices

Appendix A

Derivation of deterministic and stochastic updates

A.1 Linear system (2.12) for $\boldsymbol{u}_{0,1}^l$, $\boldsymbol{u}_{0,2}^l$, and $\boldsymbol{\lambda}_0^l$ updates

We assume all unknowns are fixed except the vector $\boldsymbol{u}_{0,1}^l$. Taking the derivative of π in (2.4) with respect to $\boldsymbol{u}_{0,1}^l$, $l = 1, \ldots, r$, leads to the following linear system

$$\sum_{l'=1}^{r} \left(\mathbb{E}\left[\phi_1^l \phi_2^l \boldsymbol{K}_1 \phi_1^{l'} \phi_2^{l'}\right] \boldsymbol{u}_{0,1}^{l'} - \mathbb{E}\left[\phi_1^l \phi_2^l \phi_1^{l'} \phi_2^{l'}\right] \boldsymbol{C}_1 \boldsymbol{\lambda}_0^{l'} \right) = \mathbb{E}\left[\phi_1^l \phi_2^l \boldsymbol{f}_1\right], \quad l = 1, \ldots, r. \tag{A.1}$$

Similarly, the $\boldsymbol{u}_{0,2}^l$ updates are performed through

$$\sum_{l'=1}^{r} \left(\mathbb{E}\left[\phi_1^l \phi_2^l \boldsymbol{K}_2 \phi_1^{l'} \phi_2^{l'}\right] \boldsymbol{u}_{0,2}^{l'} + \mathbb{E}\left[\phi_1^l \phi_2^l \phi_1^{l'} \phi_2^{l'}\right] \boldsymbol{C}_2 \boldsymbol{\lambda}_0^{l'} \right) = \mathbb{E}\left[\phi_1^l \phi_2^l \boldsymbol{f}_2\right], \quad l = 1, \ldots, r. \tag{A.2}$$

The derivative of π with respect to deterministic vector $\boldsymbol{\lambda}_0^l$ of the Lagrange multipliers satisfies

$$\sum_{l'=1}^{r} \left(\mathbb{E}\left[\phi_1^l \phi_2^l \phi_1^{l'} \phi_2^{l'}\right] \boldsymbol{C}_2^T \boldsymbol{u}_{0,2}^{l'} - \mathbb{E}\left[\phi_1^l \phi_2^l \phi_1^{l'} \phi_2^{l'}\right] \boldsymbol{C}_1^T \boldsymbol{u}_{0,1}^{l'} \right) = \boldsymbol{0}, \quad l = 1, \ldots, r. \tag{A.3}$$

Putting (A.1), (A.2), and (A.3) together results in the linear system of (2.12) for $\boldsymbol{u}_{0,1}^l$, $\boldsymbol{u}_{0,2}^l$, and $\boldsymbol{\lambda}_0^l$ updates.

A.2 Linear system (2.17) for $\phi_1^l(\boldsymbol{\xi}_1)$ updates

We first remind that in $\phi_1^l(\boldsymbol{\xi}_1)$ updates the vectors $\boldsymbol{u}_{0,1}^l$, $\boldsymbol{u}_{0,2}^l$, and $\boldsymbol{\lambda}_0^l$ are fixed at their current values obtained from (2.12). As the solution to (2.12) satisfies the constraints in (A.3), the term $\mathbb{E}\left[\boldsymbol{\lambda}^T \left(\boldsymbol{C}_2^T \boldsymbol{u}_2 - \boldsymbol{C}_1^T \boldsymbol{u}_1\right)\right]$ in (2.4) corresponding to the the interface condition vanishes for any $\{\phi_1^l(\boldsymbol{\xi}_1)\}_{l=1}^r$ and $\{\phi_2^l(\boldsymbol{\xi}_2)\}_{l=1}^r$. There-

fore, the $\phi_1^l(\boldsymbol{\xi}_1)$ updates correspond to minimization of π. More precisely, plugging the separated representation (2.9) into condition (2.14), we arrive at

$$\begin{aligned}
&\int\int\left(\sum_{l'=1}^{r}\boldsymbol{u}_{0,1}^{l^T}\phi_2^l(\boldsymbol{\xi}_2)\boldsymbol{K}_1(\boldsymbol{\xi}_1)\boldsymbol{u}_{0,1}^{l'}\phi_1^{l'}(\boldsymbol{\xi}_1)\phi_2^{l'}(\boldsymbol{\xi}_2)\right)\delta\phi_1^l(\boldsymbol{\xi}_1)\rho_1(\boldsymbol{\xi}_1)\mathrm{d}\boldsymbol{\xi}_1\rho_2(\boldsymbol{\xi}_2)\mathrm{d}\boldsymbol{\xi}_2\\
&+\int\int\left(\sum_{l'=1}^{r}\boldsymbol{u}_{0,2}^{l^T}\phi_2^l(\boldsymbol{\xi}_2)\boldsymbol{K}_2(\boldsymbol{\xi}_2)\boldsymbol{u}_{0,2}^{l'}\phi_1^{l'}(\boldsymbol{\xi}_1)\phi_2^{l'}(\boldsymbol{\xi}_2)\right)\delta\phi_1^l(\boldsymbol{\xi}_1)\rho_1(\boldsymbol{\xi}_1)\mathrm{d}\boldsymbol{\xi}_1\rho_2(\boldsymbol{\xi}_2)\mathrm{d}\boldsymbol{\xi}_2\\
&=\int\int\left(\boldsymbol{u}_{0,1}^{l^T}\boldsymbol{f}_1\phi_2^l(\boldsymbol{\xi}_2)+\boldsymbol{u}_{0,2}^{l^T}\boldsymbol{f}_2\phi_2^l(\boldsymbol{\xi}_2)\right)\delta\phi_1^l(\boldsymbol{\xi}_1)\rho_1(\boldsymbol{\xi}_1)\mathrm{d}\boldsymbol{\xi}_1\rho_2(\boldsymbol{\xi}_2)\mathrm{d}\boldsymbol{\xi}_2,\quad\forall\,\delta\phi_1^l(\boldsymbol{\xi}_1),
\end{aligned}\tag{A.4}$$

where $\rho_1(\boldsymbol{\xi}_1)$ and $\rho_2(\boldsymbol{\xi}_2)$ are the joint probability density functions of $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$, respectively. One can rewrite (A.4) as

$$\begin{aligned}
&\int\left(\sum_{l'=1}^{r}\boldsymbol{u}_{0,1}^{l^T}\boldsymbol{K}_1(\boldsymbol{\xi}_1)\boldsymbol{u}_{0,1}^{l'}\phi_1^{l'}(\boldsymbol{\xi}_1)\mathbb{E}_{\boldsymbol{\xi}_2}\left[\phi_2^l\phi_2^{l'}\right]\right)\delta\phi_1^l(\boldsymbol{\xi}_1)\rho_1(\boldsymbol{\xi}_1)\mathrm{d}\boldsymbol{\xi}_1\\
&+\int\left(\sum_{l'=1}^{r}\boldsymbol{u}_{0,2}^{l^T}\mathbb{E}_{\boldsymbol{\xi}_2}\left[\boldsymbol{K}_2\phi_2^l\phi_2^{l'}\right]\boldsymbol{u}_{0,2}^{l'}\phi_1^{l'}(\boldsymbol{\xi}_1)\right)\delta\phi_1^l(\boldsymbol{\xi}_1)\rho_1(\boldsymbol{\xi}_1)\mathrm{d}\boldsymbol{\xi}_1\\
&=\int\left(\boldsymbol{u}_{0,1}^{l^T}\boldsymbol{f}_1\mathbb{E}_{\boldsymbol{\xi}_2}\left[\phi_2^l\right]+\boldsymbol{u}_{0,2}^{l^T}\boldsymbol{f}_2\mathbb{E}_{\boldsymbol{\xi}_2}\left[\phi_2^l\right]\right)\delta\phi_1^l(\boldsymbol{\xi}_1)\rho_1(\boldsymbol{\xi}_1)\mathrm{d}\boldsymbol{\xi}_1,\quad\forall\,\delta\phi_1^l(\boldsymbol{\xi}_1),
\end{aligned}$$

which, if the integrands are continuous, is equivalent to

$$\begin{aligned}
&\sum_{l'=1}^{r}\boldsymbol{u}_{0,1}^{l^T}\boldsymbol{K}_1(\boldsymbol{\xi}_1)\boldsymbol{u}_{0,1}^{l'}\phi_1^{l'}(\boldsymbol{\xi}_1)\mathbb{E}_{\boldsymbol{\xi}_2}\left[\phi_2^l\phi_2^{l'}\right]+\sum_{l'=1}^{r}\boldsymbol{u}_{0,2}^{l^T}\mathbb{E}_{\boldsymbol{\xi}_2}\left[\boldsymbol{K}_2\phi_2^l\phi_2^{l'}\right]\boldsymbol{u}_{0,2}^{l'}\phi_1^{l'}(\boldsymbol{\xi}_1)\\
&=\boldsymbol{u}_{0,1}^{l^T}\boldsymbol{f}_1\mathbb{E}_{\boldsymbol{\xi}_2}\left[\phi_2^l\right]+\boldsymbol{u}_{0,2}^{l^T}\boldsymbol{f}_2\mathbb{E}_{\boldsymbol{\xi}_2}\left[\phi_2^l\right],\qquad l=1,\ldots,r.
\end{aligned}\tag{A.5}$$

It is straightforward to check that (2.15) – with its components given in (2.16) – is the matrix representation of (A.5). A similar approach can be used to derive (2.17).

Appendix B

Wiener chaos coefficients of a lognormal random field

Let $\kappa(\boldsymbol{x}, \boldsymbol{\xi}) = \exp(G(\boldsymbol{x}, \boldsymbol{\xi}))$ be a lognormal random field, where $G(\boldsymbol{x}, \boldsymbol{\xi})$ is a Gaussian random field given by the Karhunen-Loève expansion $G = \bar{G} + \sum_{j=1}^{d} \sqrt{\tau_j} g_j(\boldsymbol{x}) \xi_j$. Here, $\bar{G}$ is the mean of G, $\{\tau_j\}_{j=1}^{d_i}$ and $\{g_j(\boldsymbol{x})\}_{j=1}^{d_i}$ are, respectively, d largest eigenvalues and the corresponding eigenfunctions of the covariance function of G, and $\{\xi_j\}_{j=1}^{d}$ are i.i.d. normal Gaussian random variables.

Following [246], the coefficients $\kappa_{\boldsymbol{i}}(\boldsymbol{x})$ in the Hermite polynomial chaos expansion $\kappa(\boldsymbol{x}, \boldsymbol{\xi}) \approx \sum_{\boldsymbol{j} \in \mathscr{I}_{d,p}} \kappa_{\boldsymbol{j}}(\boldsymbol{x}) \psi_{\boldsymbol{j}}(\boldsymbol{\xi})$ can be computed by

$$\kappa_{\boldsymbol{i}}(\boldsymbol{x}) = \frac{\bar{\kappa}}{\sqrt{\boldsymbol{i}!}} \prod_{j=1}^{d} \left[\sqrt{\tau_j} g_j(\boldsymbol{x})\right]^{i_j}, \tag{B.1}$$

where $\bar{\kappa} = \exp\left[\bar{G} + \frac{\mathrm{var}[G]}{2}\right]$, and $\sqrt{\boldsymbol{i}!} = \prod_{j=1}^{d} i_j!$.

Appendix C

Matrix determinant update formula

Let $\boldsymbol{A}$ be an invertible matrix of size $P \times P$, and let $\boldsymbol{b}$ and $\boldsymbol{a}$ be two vectors of size P. Then

$$\frac{|\boldsymbol{A} \pm \boldsymbol{b}\boldsymbol{a}^T|}{|\boldsymbol{A}|} = 1 \pm \boldsymbol{a}^T \boldsymbol{A}^{-1} \boldsymbol{b}. \tag{C.1}$$

Proof. When $\boldsymbol{A} = \boldsymbol{I}$, we have

$$\begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{a}^T & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{I} \pm \boldsymbol{b}\boldsymbol{a}^T & \boldsymbol{b} \\ \boldsymbol{0} & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \mp \boldsymbol{a}^T & 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{b} \\ \boldsymbol{0} & \boldsymbol{I} \pm \boldsymbol{a}^T \boldsymbol{b} \end{bmatrix}.$$

Taking the determinant of both sides results in $|\boldsymbol{I} \pm \boldsymbol{b}\boldsymbol{a}^T| = 1 \pm \boldsymbol{a}^T \boldsymbol{b}$. Finally, one may write $|\boldsymbol{A} \pm \boldsymbol{b}\boldsymbol{a}^T| = |\boldsymbol{A}(\boldsymbol{I} \pm \boldsymbol{A}^{-1}\boldsymbol{y}\boldsymbol{a}^T)| = |\boldsymbol{A}|(1 \pm \boldsymbol{a}^T \boldsymbol{A}^{-1} \boldsymbol{b})$. □

Appendix D

Matrix trace update formula

Let $\boldsymbol{A}$ be an invertible matrix of size $P \times P$, and let $\boldsymbol{b}$ and $\boldsymbol{a}$ be two vectors of size P. Additionally, assume $1 \pm \boldsymbol{a}^T A^{-1} \boldsymbol{b} \neq 0$. Then

$$\text{Tr}\left((\boldsymbol{A} \pm \boldsymbol{b}\boldsymbol{a}^T)^{-1}\right) = \text{Tr}\left(\boldsymbol{A}^{-1}\right) \mp \text{Tr}\left(\frac{\boldsymbol{A}^{-1}\boldsymbol{b}\boldsymbol{a}^T\boldsymbol{A}^{-1}}{1 \pm \boldsymbol{a}^T A^{-1}\boldsymbol{b}}\right). \tag{D.1}$$

Proof. Sherman-Morrison rank-one update formula states

$$(\boldsymbol{A} \pm \boldsymbol{b}\boldsymbol{a}^T)^{-1} = \boldsymbol{A}^{-1} \mp \left(\frac{\boldsymbol{A}^{-1}\boldsymbol{b}\boldsymbol{a}^T\boldsymbol{A}^{-1}}{1 \pm \boldsymbol{a}^T A^{-1}\boldsymbol{b}}\right), \tag{D.2}$$

which can be easily verified by substituting (D.2) in identity $(\boldsymbol{A} \pm \boldsymbol{b}\boldsymbol{a}^T)(\boldsymbol{A} \pm \boldsymbol{b}\boldsymbol{a}^T)^{-1} = \boldsymbol{I}$. Taking the trace of both sides of (D.2) completes the proof. □

Bibliography

[1] A. A. FALCÓ AND A. NOUY, Proper generalized decomposition for nonlinear convex problems in tensor banach spaces, Numerische Mathematik, 121 (2012), pp. 503–530.

[2] O. B. W. A. TOSELLI, Domain Decomposition Methods Algorithms and Theory, Springer, Berlin, 2005.

[3] E. ACAR, S. A. ÇAMTEPE, M. S. KRISHNAMOORTHY, AND B. YENER, Collective sampling and analysis of high order tensors for chatroom communications, in ISI 2005: Proceedings of the IEEE International Conference on Intelligence and Security Informatics, 2006, pp. 213–224.

[4] A. AMMAR, B. MOKDAD, F. CHINESTA, AND R. KEUNINGS, A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modeling of complex fluids, Journal of Non-Newtonian Fluid Mechanics, 139 (2006), pp. 153 – 176.

[5] C. M. ANDERSON-COOK, C. M. BORROR, AND D. C. MONTGOMERY, Response surface design evaluation and comparison, Journal of Statistical Planning and Inference, 139 (2009), pp. 629 – 641.

[6] V. ARAVINDAN, J. GNANARAJ, Y.-S. LEE, AND S. MADHAVI, $LiMnPO_4$ - a next generation cathode material for lithium-ion batteries, J. Mater. Chem. A, 1 (2013), pp. 3518–3539.

[7] M. ARNST, R. GHANEM, E. PHIPPS, AND J. RED-HORSE, Reduced chaos expansions with random coefficientsin reduced-dimensional stochastic modeling of coupled problems, International Journal for Numerical Methods in Engineering, 97 (2014), pp. 352–376.

[8] A. B. ATASHGAH AND A. SEIFI, Optimal design of multi-response experiments using semi-definite programming, Optimization and Engineering, 10 (2008), pp. 75–90.

[9] A. ATKINSON, A. DONEV, AND R. TOBIAS, Optimum experimental designs, with SAS, Oxford University Press, 2007.

[10] A. C. ATKINSON AND A. N. DONEV, The construction of exact D-optimum experimental designs with application to blocking response surface designs, Biometrika, 76 (1989), pp. 515–526.

[11] I. BABUŠKA, F. NOBILE, AND R. TEMPONE, A stochastic collocation method for elliptic partial differential equations with random input data, Tech. Rep. 05-47, The Institute for Computational Engineering and Sciences (ICES), University of Texas, Austin, 2005. http://www.ices.utexas.edu/research/reports/2005.php.

[12] I. BABUŠKA, F. NOBILE, AND R. TEMPONE, A stochastic collocation method for elliptic partial differential equations with random input data, SIAM Journal on Numerical Analysis, 45 (2007), pp. 1005–1034.

[13] A. BANASZUK, V. A. FONOBEROV, T. A. FREWEN, M. KOBILAROV, G. MATHEW, I. MEZIC, A. PINTO, T. SAHAI, H. SANE, A. SPERANZON, AND A. SURANA, Scalable approach to uncertainty quantification and robust design of interconnected dynamical systems, Annual Reviews in Control, 35 (2011), pp. 77 – 98.

[14] P. BARALDI, F. CADINI, F. MANGILI, AND E. ZIO, Model-based and data-driven prognostics under different available information, Probabilistic Engineering Mechanics, 32 (2013), pp. 66 – 79.

[15] A. J. BARD AND L. R. FAULKNER, Electrochemical Methods: Fundamentals and Applications, Wiley, New York, 1980.

[16] E. V. BERG AND M. P. FRIEDLANDER, SPGL1: A solver for large-scale sparse reconstruction, June 2007. Available from http://www.cs.ubc.ca/labs/scl/spgl1.

[17] ——, Probing the Pareto frontier for basis pursuit solutions, SIAM Journal on Scientific Computing, 31 (2008), pp. 890–912.

[18] M. P. F. BERGER AND W. K. WONG, An introduction to optimal designs for social and biomedical research, John Wiley & Sons, Ltd, 2009.

[19] D. BERTSIMAS AND A. THIELE, Robust and Data-Driven Optimization: Modern Decision Making Under Uncertainty, 2014, ch. 5, pp. 95–122.

[20] G. BEYLKIN AND M. MOHLENKAMP, Numerical operator calculus in higher dimensions, Proceedings of the National Academy of Science, 99 (2002), pp. 10246–10251.

[21] ——, Algorithms for numerical analysis in high dimensions, SIAM Journal on Scientific Computing, 26 (2005), pp. 2133–2159.

[22] E. BICKEL AND R. BRATVOLD, From uncertainty quantification to decision making in the oil and gas industry, Energy, Exploration & Exploitation, 26 (2008), pp. 311–325.

[23] M. BIERI AND C. SCHWAB, Sparse high order fem for elliptic sPDEs, Computer Methods in Applied Mechanics and Engineering, 198 (2009), pp. 1149–1170.

[24] G. BLATMAN, B. SUDRET, AND M. BERVEILLER, Quasi random numbers in stochastic finite element analysis, Mcanique & Industries, 8 (2007), pp. 289–297.

[25] G. BOX, J. HUNTER, AND W. HUNTER, Statistics for Experimenters: Design, Innovation, and Discovery, Wiley, 2nd ed., 2005.

[26] P. BRO, Parafac. tutorial and applications, Chemometrics and Intelligent Laboratory Systems, 38 (1997), pp. 149–171.

[27] A. BROUDISCOU, R. LEARDI, AND R. PHAN-TAN-LUU, Genetic algorithm as a tool for selection of D-optimal design, Chemometrics and Intelligent Laboratory Systems, 35 (1996), pp. 105 – 116.

[28] L. BROWN, I. OLKIN, J. SACKS, AND H. P. WYNN (EDS.), Jack Kiefer. Collected Papers III. Design of Experiments, New York: Springer, 1985.

[29] A. BRUCKSTEIN, D. DONOHO, AND M. ELAD, From sparse solutions of systems of equations to sparse modeling of signals and images, SIAM Review, 51 (2009), pp. 34–81.

[30] E. BURNAEV, I. PANIN, AND B. SUDRET, Conformal and Probabilistic Prediction with Applications: 5th International Symposium, COPA 2016, Madrid, Spain, April 20-22, 2016, Proceedings, Springer International Publishing, Cham, 2016, ch. Effective Design for Sobol Indices Estimation Based on Polynomial Chaos Expansions, pp. 165–184.

[31] N. A. BUTLER, Optimal and orthogonal Latin hypercube designs for computer experiments, Biometrika, 88 (2001), pp. 847–857.

[32] R. E. CAFLISCH, Monte carlo and quasi-monte carlo methods, Acta numerica, 7 (1998), pp. 1–49.

[33] L. CAI AND R. E. WHITE, Reduction of model order based on proper orthogonal decomposition for lithium-ion battery simulations, J. Electrochem. Soc., 156 (2009), pp. A154–A161.

[34] R. CAMERON AND W. MARTIN, The orthogonal development of nonlinear functionals in series of Fourier-Hermite functionals, Ann. Math., 48 (1947), pp. 385–392.

[35] E. CANDES AND Y. PLAN, A probabilistic and ripless theory of compressed sensing, Information Theory, IEEE Transactions on, 57 (2011), pp. 7235–7254.

[36] E. CANDES AND T. TAO, Near optimal signal recovery from random projections: Universal encoding strategies?, IEEE Transactions on information theory, 52 (2006), pp. 5406–5425.

[37] T. F. CHAN AND T. P. MATHEW, Domain decomposition algorithms, Third International Symposium on Domain Decomposition Methods for Partial Differential Equations, T. F. Chan, R. Glowinski, J. Periaux, and O. B. Widlund, eds., SIAM, Philadelphia, (1990).

[38] D. C. CHARMPIS AND M. PAPADRAKAKIS, Enhancing the performance of the feti method with preconditioning techniques implemented on clusters of networked computers, Computational Mechanics, 30 (2002), pp. 12–28.

[39] J. CHEN, Recent progress in advanced materials for lithium ion batteries, Materials, 6 (2013), pp. 156–183.

[40] Y.-H. CHEN, C.-W. WANG, G. LIU, X.-Y. SONG, V. S. BATTAGLIA, AND A. M. SASTRYA, Selection of conductive additives in li-ion battery cathodes. a numerical study, J. Electrochem. Soc., 154 (2007), pp. A978–A986.

[41] M. CHEVREUIL, A. NOUY, AND E. SAFATLY, A multiscale method with patch for the solution of stochastic partial differential equations with localized uncertainties, Computer Methods in Applied Mechanics and Engineering, 255 (2013), pp. 255–274.

[42] S. K. CHOI, R. CANFIELD, R. GRANDHI, AND C. PETTIT, Polynomial Chaos Expansion with Latin Hypercube Sampling for Estimating Response Variability, AIAA Journal, 42 (2004), pp. 1191–1198.

[43] C.-S. CHOU, C.-H. TSOU, AND C.-I. WANG, Preparation of graphite/nano-powder composite particles and applicability as carbon anode material in a lithium ion battery, Advanced Powder Technology, 19 (2008), pp. 383–396.

[44] D.-W. CHUNG, P. R. SHEARING, N. P. BRANDON, S. J. HARRIS, AND R. E. GARCIA, Particle size polydispersity in li-ion batteries, J. Electrochem. Soc., 161 (2014), pp. A422–A430.

[45] A. COHEN, M. A. DAVENPORT, AND D. LEVIATAN, On the stability and accuracy of least squares approximations, Foundations of Computational Mathematics, 13 (2013), pp. 819–834.

[46] P. G. CONSTANTINE, M. ELDRED, AND E. PHIPPS, Sparse pseudospectral approximation method, Computer Methods in Applied Mechanics and Engineering, 229 (2012), pp. 1–12.

[47] P. G. CONSTANTINE, M. S. ELDRED, AND E. T. PHIPPS, Sparse pseudospectral approximation method, Computer Methods in Applied Mechanics and Engineering, 229232 (2012), pp. 1 – 12.

[48] R. D. COOK AND C. J. NACHTSHEIM, A comparison of algorithms for constructing exact D-optimal designs, Technometrics, 22 (1980), pp. 315–324.

[49] R. COTTEREAU, D. CLOUTEAU, H. B. DHIA, AND C. ZACCARDI, A stochastic-deterministic coupling method for continuum mechanics, Computer Methods in Applied Mechanics and Engineering, 200 (2011), pp. 3280–3288.

[50] M. DAIGLE AND C. KULKARNI, Electrochemistry-based battery modeling for prognostics, in Annual Conference of the Prognostics and Health Management Society, October 2013, pp. 249–261.

[51] M. DAIGLE, A. SAXENA, AND K. GOEBEL, An efficient deterministic approach to model-based prediction uncertainty estimation, in: Annual Conference of the Prognostics and Health Management Society, (2012), pp. 326–335.

[52] R. DARLING AND J. NEWMAN, Modeling a porous intercalation electrode with two characteristic particle sizes, J. Electrochem. Soc., 144 (1997), pp. 4201–4208.

[53] H. DETTE AND Y. GRIGORIEV, E-optimal designs for second-order response surface models, Ann. Statist., 42 (2014), pp. 1635–1656.

[54] J. DICK AND F. PILLICHSHAMMER, Digital Nets and Sequences: Discrepancy Theory and Quasi-Monte Carlo Integration, Cambridge University Press, New York, NY, USA, 2010.

[55] C. R. DOHRMANN, A preconditioner for substructuring based on constrained energy minimization, SIAM J. Sci. Comput., 25 (2003), pp. 246–258.

[56] D. DONOHO, Compressed sensing, IEEE Transactions on information theory, 52 (2006), pp. 1289–1306.

[57] A. DOOSTAN AND G. IACCARINO, A least-squares approximation of partial differential equations with high-dimensional random inputs, Journal of Computational Physics, 228 (2009), pp. 4332–4345.

[58] A. DOOSTAN, G. IACCARINO, AND N. ETEMADI, A least-squares approximation of high-dimensional uncertain systems, Tech. Rep. Annual Research Brief, Center for Turbulence Research, Stanford University, 2007.

[59] A. DOOSTAN AND H. OWHADI, A non-adapted sparse approximation of PDEs with stochastic inputs, Journal of Computational Physics, 230 (2011), pp. 3015–3034.

[60] A. DOOSTAN, A. VALIDI, AND G. IACCARINO, Non-intrusive low-rank separated approximation of high-dimensional stochastic models, Computer Methods in Applied Mechanics and Engineering, (2013). in press.

[61] M. DOYLE AND Y. FUENTES, Computer simulations of a lithium-ion polymer battery and implications for higher capacity next-generation battery designs, Journal of The Electrochemical Society, 150 (2003), pp. A706–A713.

[62] M. DOYLE, T. FULLER, AND J. NEWMAN, Modeling of galvanostatic charge and discharge of the lithium/polymer/insertion cell, Journal of the Electrochemical Society, 140 (1993), pp. 1526–1533.

[63] M. DOYLE, J. NEWMAN, A. S. GOZDZ, C. N. SCHMUTZ, AND J. TARASCON, Comparison of modeling predictions with experimental data from plastic lithium ion cells, J. Electrochem. Soc., 143 (1996), pp. 1890–1903.

[64] W. DU, N. XUE, W. SHYY, AND J. R. R. A. MARTINS, A surrogate-based multi-scale model for mass transport and electrochemical kinetics in Lithium-ion battery electrodes, J. Electrochem. Soc., 161 (2014), pp. E3086–E3096.

[65] W. DUA, A. GUPTA, X. ZHANG, A. M. SASTRY, AND W. SHYY, Effect of cycling rate, particle size and transport properties on lithium-ion cathode performance, International Journal of Heat and Mass Transfer, 53 (2010), pp. 3552–3561.

[66] T. DUBESHTER, P. K. SINHA, A. SAKARS, G. W. FLY, , AND J. JORNE, Measurement of tortuosity and porosity of porous battery electrodes, J. Electrochem. Soc., 161 (2014), pp. A599–A605.

[67] R. DUVIGNEAU, M. MARTINELLI, AND P. CHANDRASHEKARAPPA, Uncertainty Quantification for Robust Design, John Wiley & Sons, Inc., 2013, pp. 405–424.

[68] O. DYKSTRA, The augmentation of experimental data to maximize $|x'x|$, Technometrics, 13 (1971), pp. 682–688.

[69] M. EBNER, D.-W. CHUNG, R. E. GARCIA, AND V. WOOD, Tortuosity anisotropy in lithium-ion battery electrodes, Adv. Energy Mater., 4 (2014), p. 1301278.

[70] M. S. ELDRED, Recent advances in non-intrusive polynomial chaos and stochastic collocation methods for uncertainty analysis and design, in 50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, Palm Springs, CA, 2009.

[71] K. ENTACHER, Discrepancy estimates based on haar functions, Mathematics and Computers in Simulation, 55 (2001), pp. 49 – 57. The Second {IMACS} Seminar on Monte Carlo Methods.

[72] V. ETACHERI, R. MAROM, R. ELAZARI, G. SALITRA, AND D. AURBACH, Challenges in the development of advanced li-ion batteries: a review, Energy Environ. Sci., 4 (2011), p. 3243.

[73] D. FALLER, U. KLINGMLLER, AND J. TIMMER, Simulation methods for optimal experimental design in systems biology, Simulation, (2003), p. 2003.

[74] K. FANG, D. K. J. LIN, P. WINKER, AND Y. ZHANG, Uniform design: Theory and application, Technometrics, 42 (2000), pp. 237–248.

[75] K. T. FANG, The Uniform Design: Application of Number-Theoretic Methods in Experimental Design, Acta Mathematicae Applicatae Sinica 3(4) December 197, 3 (1980), pp. 363–372.

[76] K. T. FANG AND H. QIN, A note on construction of nearly uniform designs with large number of runs, Statistics & Probability Letters, 61 (2003), pp. 215 – 224.

[77] K. T. FANG, W.-C. SHIU, AND J.-X. PAN, Uniform designs based on Latin squares, Statistica Sinica, 9 (1999), pp. 905–912.

[78] K. T. FANG AND Y. WANG, Number-Theoretic Methods in Statistics, Chapman & Hall, London, 1994.

[79] C. FARHAT, P. S. CHEN, AND J. MANDEL, A scalable lagrange multiplier based domain decomposition method for time-dependent problems, Int. J. Numer. Meth. Eng., 38 (1995), pp. 3831–3853.

[80] C. FARHAT, M. LESOINNE, P. LETALLEC, K. PIERSON, AND D. RIXEN, Feti-dp: a dual-primal unified feti method–part i: a faster alternative to the two-level feti method, Int. J. Num. Meth. Eng., 50 (2001), pp. 1523–1544.

[81] C. FARHAT AND J. MANDEL, The two-level feti method for static and dynamic plate problemspart i: An optimal iterative solver for biharmonic systems, Comput. Methods in App. Mech. Eng., 155 (1998), pp. 129–152.

[82] C. FARHAT AND F. ROUX, A method of finite element tearing and interconnecting and its parallel solution algorithm, International Journal for Numerical Methods in Engineering, 32 (1991), pp. 1205–1227.

[83] V. FEDOROV AND P. HACKL, Model-Oriented Design of Experiments, Springer-Verlag New York, 1997.

[84] V. FEDOROV, W. STUDDEN, AND E. KLIMKO, Theory of Optimal Experiments, Academic Press Inc, 1972.

[85] A. FERRY, M. M. DOEFF, AND L. C. D. JONGHE, Transport property and raman spectroscopic studies of the polymer electrolyte system $p(eo)_n - natfsi$, J. Electrochem. Soc., 145 (1998), pp. 1586–1592.

[86] J. FOO AND G. KARNIADAKIS, Multi-element probabilistic collocation method in high dimensions, J. Comput. Phys., 229 (2010), pp. 1536–1557.

[87] M. FORACCHIA, A. C. HOOKER, P. VICINI, AND A. RUGGERI, POPED, a software for optimal experiment design in population kinetics, Computer Methods and Programs in Biomedicine, 74 (2004).

[88] J. C. FORMAN, S. BASHASH, J. L. STEIN, AND H. K. FATHY, Reduction of an electrochemistry-based li-ion battery model via quasi-linearization and pade approximation, J. Electrochem. Soc., 158 (2011), pp. A93–A101.

[89] L. FRANCA AND A. MACEDO, A two-level finite element method and its application to the helmholtz equation, Int. J. Numer. Meth. Eng., 43 (1998), pp. 23–32.

[90] T. F. FULLER, M. DOYLE, AND J. NEWMAN, Simulation and optimization of the dual lithium ion insertion cell, Journal of The Electrochemical Society, 141 (1994), pp. 1–10.

[91] R. FURUKAWA, H. KAWASAKI, K. IKEUCHI, AND M. SAKAUCHI, Appearance based object modeling using texture database: acquisition, compression and rendering, in EGRW 02: Proceedings of the 13th Eurographics workshop on Rendering, Aire-la-Ville, Switzerland, Switzerland, 2001, pp. 257–266.

[92] J. GALVIS AND M. SARKIS, Feti and bdd preconditioners for stokesmortardarcy systems, Comm. App. Math. and Comp. Sci., 5 (2010), pp. 1–30.

[93] Z. GAO AND J. S. HESTHAVEN, On ANOVA expansions and strategies for choosing the anchor point, Tech. Rep. 2010-17, Scientific Computing Group, Brown University, Providence, RI, USA, 2010.

[94] M. GERADIN, D. COULON, AND J. DELSEMME, Parallelization of the samcef finite element software through domain decomposition and feti algorithm, International Journal of Supercomputer Applications, 11 (1997), pp. 286–298.

[95] R. GHANEM AND P. SPANOS, Stochastic Finite Elements: A Spectral Approach, Springer Verlag, 1991.

[96] D. GHOSH, P. AVERY, AND C. FARHAT, A feti-preconditioned congugate gradient method for large-scale stochastic finite element problems, Int. J. Numer. Meth. Engng, 80 (2009), pp. 914–931.

[97] D. GHOSH AND C. FARHAT, Strain and stress computations in stochastic finite element methods, International Journal for Numerical Methods in Engineering, 74 (2008), pp. 1219–1239.

[98] W. R. GILKS, Markov Chain Monte Carlo In Practice, Chapman and Hall/CRC, 1999.

[99] S. GOLMON, K. MAUTE, AND M. DUNN, Numerical modeling of electrochemical-mechanical interactions in lithium polymer batteries, Computers and Structures, 87 (2009), pp. 1567–1579.

[100] S. GOLMON, K. MAUTE, AND M. L. DUNN, A pomegranate-inspired nanoscale design for large-volume-change lithium battery anodes, International Journal for Numerical Methods in Engineering, 92 (2012), pp. 475–494.

[101] P. GOOS AND B. JONES, Optimal Design of Experiments: A Case Study Approach, Wiley, 2011.

[102] P. GOSSELET AND C. REY, Non-overlapping domain decomposition methods in structural mechanics, Archives of Computational Methods in Engineering, 13 (2006), pp. 515–572.

[103] M. GRABE, Measurement Uncertainties in Science and Technology, Springer, 2005.

[104] W. HACKBUSCH, Tensor spaces and numerical tensor calculus, vol. 42, Springer, 2012.

[105] M. HADIGOL, K. MAUTE, AND A. DOOSTAN, On uncertainty quantification of lithium-ion batteries: Application to an $Lic_6/LiCoO_2$ cell, Journal of Power Sources, 300 (2015), pp. 507–524.

[106] L. M. HAINES, The application of the annealing algorithm to the construction of exact optimal designs for linear-regression models, Technometrics, 29 (1987), pp. 439–447.

[107] J. HAMPTON AND A. DOOSTAN, Coherence motivated sampling and convergence analysis of least squares polynomial chaos regression, Computer Methods in Applied Mechanics and Engineering, 290 (2015), pp. 73 – 97.

[108] ——, Compressive sampling of polynomial chaos expansions: Convergence analysis and sampling strategies, Journal of Computational Physics, 280 (2015), pp. 363 – 386.

[109] R. HARDIN AND N. SLOANE, A new approach to the construction of optimal designs, Journal of Statistical Planning and Inference, 37 (1993), pp. 339 – 369.

[110] R. H. HARDIN AND N. J. A. SLOANE, Gosset: A general-purpose program for designing experiments. http://www.NeilSloane.com/gosset.

[111] F. HITCHCOCK, The expression of a tensor or a polyadic as a sum of products, Journal of Mathematics and Physics, 6 (1927), pp. 164–189.

[112] R. HOOKE AND T. A. JEEVES, Direct search solution of numerical and statistical problems, J. ACM, 8 (1961), pp. 212–229.

[113] S. HOSDER, R. WALTERS, AND R. PEREZ, A non-intrusive polynomial chaos method for uncertainty propagation in CFD simulations, in 44th AIAA aerospace sciences meeting and exhibit, AIAA-2006-891, Reno (NV), 2006.

[114] D. HUNTINGTON AND C. LYRINTZIS, Improvements to and limitations of Latin hypercube sampling, Probabilistic Engineering Mechanics, 13 (1998), pp. 245 – 253.

[115] R. L. IMAN AND W. J. CONOVER, A distribution-free approach to inducing rank correlation among input variables, Communications in Statistics - Simulation and Computation, 11 (1982), pp. 311–334.

[116] R. L. IMAN, J. C. HELTON, AND J. E. Y. CAMPBELL, An approach to sensitivity analysis of computer models, part 1. introduction, input variable selection an preliminary variable assessment, Journal on Quality Technology, 13 (1981), pp. 174–183.

[117] W. J. W. JEROME SACKS, SUSANNAH B. SCHILLER, Designs for computer experiments, Technometrics, 31 (1989), pp. 41–47.

[118] D. M. JIE YANG, ABHYUDAY MANDAL, Optimal designs for two-level factorial experiments with binary response, Statistica Sinica, 22 (2012), pp. 885–907.

[119] M. JOHNSON, L. MOORE, AND D. YLVISAKER, Minimax and maximin distance designs, Journal of Statistical Planning and Inference, 26 (1990), pp. 131 – 148.

[120] M. E. JOHNSON AND C. J. NACHTSHEIM, Some guidelines for constructing exact d-optimal designs on convex design spaces, Technometrics, 25 (1983), pp. 271–277.

[121] R. T. JOHNSON, B. JONES, J. W. FOWLER, AND D. C. MONTGOMERY, Comparing designs for computer simulation experiments, in Simulation Conference, 2008. WSC 2008. Winter, Dec 2008, pp. 463–470.

[122] R. T. JOHNSON, D. C. MONTGOMERY, B. JONES, AND P. A. PARKER, Comparing computer experiments for fitting high-order polynomial metamodels, Journal of Quality Technology, 42 (2010), pp. 86–102.

[123] B. JONES AND P. GOOS, I-optimal versus D-optimal split-plot response surface designs, Working Papers 2012002, University of Antwerp, Faculty of Applied Economics, Jan. 2012.

[124] B. A. JONES, N. PARRISH, AND A. DOOSTAN, Postmaneuver collision probability estimation using sparse polynomial chaos expansions, Journal of Guidance, Control, and Dynamics, (2015), pp. 1–13.

[125] V. R. JOSEPH, Space-filling designs for computer experiments: A review, Quality Engineering, 28 (2016), pp. 28–35.

[126] V. R. JOSEPH, E. GUL, AND S. BA, Maximum projection designs for computer experiments, Biometrika, 102 (2015), pp. 371–380.

[127] V. R. JOSEPH AND Y. HUNG, Orthogonal-maximin Latin hypercube designs, Statistica Sinica, 18 (2008), pp. 171–186.

[128] J. R. I. JR., A. JOSHI, K. GREW, W. CHIU, A. TKACHUK, S. WANG, AND W. YUN, Nondestructive reconstruction and analysis of solid oxide fuel cell anodes using x-ray computed tomography at sub-50 nm resolution, J. Electrochem. Soc., 155 (2008), pp. B504–B508.

[129] M. J.R., Electrochemical Methods: Fundamentals and Applications, Wiley, New York, 1987.

[130] G. M. KABERA, L. M. HAINES, AND P. NDLOVU, The analytic construction of D-optimal designs for the two-variable binary logistic regression model without interaction, Statistics, 49 (2015), pp. 1169–1186.

[131] B. KHOROMSKIJ AND C. SCHWAB, Tensor-structured Galerkin approximation of parametric and stochastic elliptic PDEs, SIAM Journal on Scientific Computing, 33 (2011), pp. 364–385.

[132] J. KIEFER, Optimum experimental designs, Journal of the Royal Statistical Society. Series B (Methodological), 21 (1959), pp. 272–319.

[133] A. KLAWONN AND O. WIDLUND, Feti and neumann-neumann iterative substructuring methods: Connections and new results, Communications on Pure and Applied Mathematics, 54 (2001), pp. 57–90.

[134] M. KLEIBER, T. HIEN, AND D. TRAN, The Stochastic Finite Element Method: Basic Perturbation Technique and Computer Implementation, John Wiley & Sons, Chichester, 1992.

[135] J. P. C. KLEIJNEN, Design and Analysis of Simulation Experiments, Springer, 2008.

[136] T. G. KOLDA, Orthogonal tensor decompositions, SIAM Journal on Matrix Analysis and Applications, 23 (2001), pp. 243–255.

[137] T. G. KOLDA AND B. W. BADER, Tensor decompositions and applications, SIAM Review, 51 (2009), pp. 455–500.

[138] P. KROONENBERG AND J. LEEUW, Principal component analysis of three-mode data by means of alternating least squares algorithms, Psychometrika, 45 (1980), pp. 69–97.

[139] P. LADEVÈZE, The large time increments method for analysis of structures with nonlinear behavior described by internal variables (in french), C. R. Acad. Sci. Paris, 309 (1989), pp. 1095–1099.

[140] L. D. LATHAUWER AND J. CASTAING, Tensor-based techniques for the blind separation of ds-cdma signal, Signal Processing, 87 (2007), pp. 322–336.

[141] M. D. LEVI AND D. AURBACH, Diffusion coefficients of lithium ions during intercalation into graphite derived from the simultaneous measurements and modeling of electrochemical impedance and potentiostatic intermittent titration characteristics of thin graphite electrodes, J. Phys. Chem. B, 101 (1997), pp. 4641–4647.

[142] G. LIMENG, P. JINGYUE, L. DATONG, AND P. XIYUAN, Data-driven framework for lithium-ion battery remaining useful life estimation based on improved nonlinear degradation factor, in Electronic Measurement Instruments (ICEMI), 2013 IEEE 11th International Conference on, vol. 2, Aug 2013, pp. 1014–1020.

[143] D. LIU, W. XIE, H. LIAO, AND Y. PENG, An integrated probabilistic approach to lithium-ion battery remaining useful life estimation, Instrumentation and Measurement, IEEE Transactions on, 64 (2015), pp. 660–670.

[144] A. LOGG, K.-A. MARDAL, AND G. WELLS, Automated Solution of Differential Equations by the Finite Element Method, Springer, 2012.

[145] W. J. C. M. D. MCKAY, R. J. BECKMAN, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics, 21 (1979), pp. 239–245.

[146] S. MAHADEVAN, Uncertainty quantification for decision-making in engineered systems, in Proceedings of the International Symposium on Engineering under Uncertainty: Safety Assessment and Management (ISEUSAM - 2012), S. Chakraborty and G. Bhattacharya, eds., Springer India, 2013, pp. 97–117.

[147] C. V. MAI AND B. SUDRET, Polynomial chaos expansions for damped oscillators, in ICASP12 12th International Conference on Applications of Statistics and Probability in Civil Engineering, Vancouver, Canada, 2015.

[148] O. L. MAITRE AND O. KNIO, A stochastic particle-mesh scheme for uncertainty propagation in vortical flows, J. Comput. Phys., 226 (2007), p. 645771.

[149] ——, Spectral Methods for Uncertainty Quantification with Applications to Computational Fluid Dynamics, Springer, 2010.

[150] MALVERN INSTRUMENTS LTD, Mastersizer 2000 user manual. `http://www.malvern.com`, 2007. Accessed: 2014-07-29.

[151] A. MANDAL, W. K. WONG, AND Y. YU, Algorithmic searches for optimal designs, in Handbook of Design and Analysis of Experiments, A. Dean, M. Morris, J. Stufken, and D. Bingham, eds., Chapman and Hall/CRC, 2015, pp. 755–784.

[152] R. MARSH, S. VUKSON, S. SURAMPUDI, B. RATNAKUMAR, M. SMART, M. MANZO, AND P. DALTON, Li ion batteries for aerospace applications, Journal of Power Sources, 97–98 (2001), pp. 25–27.

[153] L. MATHELIN AND M. HUSSAINI, A stochastic collocation algorithm for uncertainty analysis, Tech. Rep. NAS 1.26:212153; NASA/CR-2003-212153, NASA Langley Research Center, 2003.

[154] B. MAUS AND G. VAN BREUKELEN, POBE: A computer program for optimal design of multi-subject blocked fMRI experiments, Journal of Statistical Software, 56 (2014), pp. 1–24.

[155] J. M. MCFARLAND, Uncertainty Analysis for Computer Simulations Through Validation and Calibration, PhD thesis, Nashville, TN, USA, 2008. AAI3378200.

[156] D. L. MCLEISH, Monte Carlo Simulation and Finance, Wiley, 1st ed., 2011.

[157] R. K. MEYER AND C. J. NACHTSHEIM, Constructing exact D-optimal experimental designs by simulated annealing, American Journal of Mathematical and Management Sciences, 8 (1988), pp. 329–359.

[158] ——, The coordinate-exchange algorithm for constructing exact optimal experimental designs, Technometrics, 37 (1995), pp. 60–69.

[159] G. MIGLIORATI AND F. NOBILE, Analysis of discrete least squares on multivariate polynomial spaces with evaluations at low-discrepancy point sets, Journal of Complexity, 31 (2015), pp. 517 – 542.

[160] G. MIGLIORATI, F. NOBILE, E. SCHWERIN, AND R. TEMPONE, Analysis of discrete L^2 projection on polynomial spaces with random evaluations, Found. Comput. Math., 14 (2014), pp. 419–456.

[161] G. MIGLIORATI, F. NOBILE, E. VON SCHWERIN, AND R. TEMPONE, Approximation of quantities of interest in stochastic pdes by the random discrete L^2 projection on polynomial spaces, SIAM Journal on Scientific Computing, 35 (2013), pp. A1440–A1460.

[162] T. J. MITCHELL, An algorithm for the construction of "D-Optimal" experimental designs, Technometrics, 16 (1974), pp. 203–210.

[163] G. MONTEPIEDRA, Application of genetic algorithms to the construction of exact D-optimal designs, Journal of Applied Statistics, 25 (1998), pp. 817–826.

[164] R. E. MOORE AND F. BIERBAUM, Methods and Applications of Interval Analysis (SIAM Studies in Applied and Numerical Mathematics) (Siam Studies in Applied Mathematics 2.), Soc for Industrial & Applied Math, 1979.

[165] M. MORRIS, Design of Experiments: An Introduction Based on Linear Models, Chapman and Hall/CRC, 2010.

[166] R. H. MYERS, D. C. MONTGOMERY, AND C. M. ANDERSON-COOK, Response Surface Methodology: Process and Product Optimization Using Designed Experiments, John Wiley, 2009.

[167] G. S. NAGARAJAN, J. W. V. ZEE, AND R. M. SPOTNITZ, A mathematical model for intercalation electrode behavior: I. effect of particle-size distribution on discharge capacity, J. Electrochem. Soc., 145 (1998), pp. 771–779.

[168] A. NARAYAN, J. D. JAKEMAN, AND T. ZHOU, A Christoffel function weighted least squares algorithm for collocatio ArXiv e-prints, (2014).

[169] G.-A. NAZRI AND G. PISTOIA, Lithium Batteries: Science and Technology, Springer, 2003.

[170] J. NEWMAN AND K. E. THOMAS-ALYEA, Electrochemical Systems, 3rd Edition, Wiley-Interscience, 2004.

[171] J. NEWMAN AND W. TIEDEMANN, Porous-electrode theory with battery applications, AIChE Journal, 21 (1975), pp. 25–41.

[172] N.-K. NGUYEN AND A. J. MILLER, A review of some exchange algorithms for constructing discrete D-optimal designs, Computational Statistics & Data Analysis, 14 (1992), pp. 489 – 498.

[173] H. NIEDERREITER, Quasi-monte carlo methods and pseudo-random numbers, Bull. Amer. Math. Soc., 84 (1978), pp. 957–1041.

[174] ——, Random Number Generation and quasi-Monte Carlo Methods, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1992.

[175] F. NOBILE, R. TEMPONE, AND C. G. WEBSTER, A sparse grid stochastic collocation method for partial differential equations with random input data, SIAM Journal on Numerical Analysis, 46 (2008), pp. 2309–2345.

[176] A. NOUY, A generalized spectral decomposition technique to solve a class of linear stochastic partial differential equations, Computer Methods in Applied Mechanics and Engineering, 196 (2007), pp. 4521–4537.

[177] ——, Generalized spectral decomposition method for solving stochastic finite element equations: Invariant subspace problem and dedicated algorithms, Computer Methods in Applied Mechanics and Engineering, 197 (2008), pp. 4718–4736.

[178] ——, Proper generalized decompositions and separated representations for the numerical solution of high dimensional stochastic problems, Archives of Computational Methods in Engineering, 17 (2010), pp. 403–434.

[179] J. NYBERG, S. UECKERT, E. A. STROEMBERG, S. HENNIG, M. O. KARLSSON, AND A. C. HOOKER, PopED: An extended, parallelized, nonlinear mixed effects models optimal design tool, Computer Methods and Programs in Biomedicine, 108 (2012).

[180] A. NYMAN, An Experimental and Theoretical Study of the Mass Transport in LithiumIon Battery Electrolytes, PhD thesis, Ph.D. Thesis, Kungliga Tekniska Hgskolan, 2011.

[181] A. B. OWEN, Latin supercube sampling for very high-dimensional simulations, ACM Trans. Model. Comput. Simul., 8 (1998), pp. 71–102.

[182] H. PANG, P. CHENG, H. YANG, J. LU, C. X. GUO, G. NING, AND C. M. LI, Template-free bottom-up synthesis of yolk-shell vanadium oxide as high performance cathode for lithium ion batteries, Chem. Commun., 49 (2013), pp. 1536–1538.

[183] K. C. PARK AND C. A. FELIPPA, A variational framework for solution method developments in structural mechanics, Journal of Applied Mechanics, 56/1 (1998), pp. 242–249.

[184] J. PENG, J. HAMPTON, AND A. DOOSTAN, A weighted l_1-minimization approach for sparse polynomial chaos expansions, Journal of Computational Physics, 267 (2014), pp. 92 – 111.

[185] T. PIAO, S.-M. PARK, C.-H. DOH, AND S.-I. MOONB, Intercalation of lithium ions into graphite electrodes studied by ac impedance measurements, Journal of The Electrochemical Society, 146 (1999), pp. 2794–2798.

[186] P. POPOV, Y. VUTOV, S. MARGENOV, AND O. ILIEV, Finite volume discretization of equations describing nonlinear diffusion in li-ion batteries, in Numerical Methods and Applications, I. Dimov, S. Dimova, and N. Kolkovska, eds., vol. 6046 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2011, pp. 338–346.

[187] L. PRONZATO, Optimal experimental design and some related control problems, Automatica, 44 (2008), pp. 303 – 325.

[188] L. PRONZATO AND W. G. MÜLLER, Design of computer experiments: space filling and beyond, Statistics and Computing, 22 (2011), pp. 681–701.

[189] F. PUKELSHEIM, Optimal Design of Experiments, Society for Industrial and Applied Mathematic, 2006.

[190] P. Z. G. QIAN, Sliced Latin hypercube designs, Journal of the American Statistical Association, 107 (2012), pp. 393–399.

[191] A. QUARTERONI AND A. VALLI, Domain decomposition methods for partial differential equations, vol. 10, Clarendon Press Oxford, 1999.

[192] V. RAMADESIGAN, K. CHEN, N. A. BURNS, V. BOOVARAGAVAN, R. D. BRAATZ, AND V. R. SUBRAMANIAN, Parameter estimation and capacity fade analysis of lithium-ion batteries using reformulated models, Journal of The Electrochemical Society, 158 (2011), pp. A1048–A1054.

[193] F. RAPETTI AND A. TOSELLI, A feti preconditioner for two dimensional edge element approximations of maxwell's equations on nonmatching grids, SIAM J. Sci. Comput., 23 (2001), pp. 92–108.

[194] J. N. REIMERS, Algorithmic improvements and pde decoupling, for the simulation of porous electrode cells, J. Electrochem. Soc., 160 (2013), pp. A811–A818.

[195] ——, Accurate and efficient treatment of foil currents in a spiral wound Li-ion cell, J. Electrochem. Soc., 161 (2014), pp. A118–A127.

[196] J. N. REIMERS, M. SHOESMITH, Y. S. LIN, AND L. O. VALOEN, Simulating high current discharges of power optimized Li-ion cells, J. Electrochem. Soc., 160 (2013), pp. A1870–A1884.

[197] D. J. RIXEN, C. FARHAT, R. TEZAUR, AND J. MANDEL, Theoretical comparison of the feti and algebraically partitioned feti methods, and performance comparisons with a direct sparse solver, Int. J. Numer. Meth. Engng., 46 (1999), pp. 501–533.

[198] J. SACKS, W. J. WELCH, T. J. MITCHELL, AND H. P. WYNN, Design and analysis of computer experiments, Statist. Sci., 4 (1989), pp. 409–423.

[199] G. SAGNOL, Computing optimal designs of multiresponse experiments reduces to second-order cone programming, Journal of Statistical Planning and Inference, 141 (2011), pp. 1684 – 1708.

[200] B. SAHA, C. QUACH, AND K. GOEBEL, Optimizing battery life for electric uavs using a bayesian framework, in Aerospace Conference, 2012 IEEE, March 2012, pp. 1–7.

[201] C. SALLABERRY, J. HELTON, AND S. HORA, Extension of Latin hypercube samples with correlated variables, Reliability Engineering & System Safety, 93 (2008), pp. 1047 – 1059. Bayesian Networks in Dependability.

[202] S. SANKARARAMAN, Significance, interpretation, and quantification of uncertainty in prognostics and remaining useful life prediction, Mechanical Systems and Signal Processing, 5253 (2015), pp. 228 – 247.

[203] S. SANKARARAMAN, M. DAIGLE, AND K. GOEBEL, Uncertainty quantification in remaining useful life prediction using first-order reliability methods, Reliability, IEEE Transactions on, 63 (2014), pp. 603–619.

[204] S. SANKARARAMAN, M. DAIGLE, A. SAXENA, AND K. GOEBEL, Analytical algorithms to quantify the uncertainty in remaining useful life prediction, in Aerospace Conference, 2013 IEEE, March 2013, pp. 1–11.

[205] S. SANKARARAMAN AND K. GOEBEL, Uncertainty quantification in remaining useful life of aerospace components using state space models and inverse form, 54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, (2013).

[206] C. SANKAVARAM, B. PATTIPATI, A. KODALI, K. PATTIPATI, M. AZAM, S. KUMAR, AND M. PECHT, Model-based and data-driven prognosis of automotive and electronic systems, in Automation Science and Engineering, 2009. CASE 2009. IEEE International Conference on, Aug 2009, pp. 96–101.

[207] S. SANTHANAGOPALAN AND R. E. WHITE, Modeling parametric uncertainty using polynomial chaos theory, ECS Trans., 3 (2007), pp. 243–256.

[208] ——, Quantifying cell-to-cell variations in lithium ion batteries, International Journal of Electrochemistry, 2012 (2012), p. Article ID 395838.

[209] T. J. SANTNER, B. J. WILLIAMS, AND W. I. NOTZ, The Design and Analysis of Computer Experiments, Springer, 2003.

[210] F. SAUVAGE, J.-M. TARASCON, AND E. BAUDRIN, In situ measurements of li ion battery electrode material conductivity: Application to li_xcoo_2 and conversion reactions, J. Phys. Chem. C, 111 (2007), pp. 9624–9630.

[211] D. SCHIAVAZZI, A. DOOSTAN, AND G. IACCARINO, Sparse multiresolution regression for uncertainty propagation, International Journal for Uncertainty Quantification, 4 (2014), pp. 303–331.

[212] P. SESHADRI, A. NARAYAN, AND S. MAHADEVAN, Optimal quadrature subsampling for least squares polynomial approximations, ArXiv e-prints, (2016). http://arXiv:1601.05470.

[213] G. SHAFER, A Mathematical Theory of Evidence, Princeton University Press, 1976.

[214] A. SHASHUA AND A. LEVIN, Linear image coding for regression and classification using the tensor-rank principle, in CVPR 2001: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001, pp. 42–49.

[215] P. SHEARING, J. GOLBERT, R. CHATER, AND N. BRANDON, 3d reconstruction of {SOFC} anodes using a focused ion beam lift-out technique, Chemical Engineering Science, 64 (2009), pp. 3928–3933.

[216] P. SHEARING, L. HOWARD, P. JORGENSEN, N. BRANDON, AND S. HARRIS, Characterization of the 3-dimensional microstructure of a graphite negative electrode from a li-ion battery, Electrochemistry Communications, 12 (2010), pp. 374–377.

[217] M. C. SHEWRY AND H. P. WYNN, Maximum entropy sampling, Journal of Applied Statistics, 14 (1987), pp. 165–170.

[218] D. SHIN, M. PONCINO, E. MACII, AND N. CHANG, A statistical model of cell-to-cell variation in li-ion batteries for system-level design, in Low Power Electronics and Design (ISLPED), 2013 IEEE International Symposium on, Sept 2013, pp. 94–99.

[219] Y. SHIN AND D. XIU, Nonadaptive quasi-optimal points selection for least squares linear regression, SIAM Journal on Scientific Computing, 38 (2016), pp. A385–A411.

[220] X.-S. SI, W. WANG, C.-H. HU, AND D.-H. ZHOU, Remaining useful life estimation a review on the statistical data driven approaches, European Journal of Operational Research, 213 (2011), pp. 1 – 14.

[221] N. SIDIROPOULOS, G. GIANNAKIS, AND R. BRO, Blind parafac receivers for ds-cdma systems, IEEE Transactions on Signal Processing, 48 (2000), pp. 810–823.

[222] G. SIKHA AND R. E. WHITE, Analytical expression for the impedance response for a lithium-ion cell, Journal of The Electrochemical Society, 155 (2008), pp. A893–A902.

[223] B. SMITH, P. BJORSTAD, AND W. GROPP, Domain decomposition, Cambridge University Press, 2004.

[224] K. SMITH, On the standard deviations of adjusted and interpolated values of an observed polynomial function and its constants and the guidance they give towards a proper choice of the distribution of observations, Biometrika, 12 (1918), pp. 1–85.

[225] K. A. SMITH, C. D. RAHN, AND C.-Y. WANG, Model order reduction of 1d diffusion systems via residue grouping, J. Dyn. Sys., Meas., Control, 130 (2008), p. 011012.

[226] S. SMOLYAK, Quadrature and interpolation formulas for tensor products of certain classes of functions, Soviet Mathematics, Doklady, 4 (1963), pp. 240–243.

[227] B. J. SMUCKER, By Design: Exchange Algorithms to Construct Exact Model-robust and Multiresponse Experimental Designs, PhD thesis, Ph.D. Thesis, Pennsylvania State University, 2010. AAI3436195.

[228] I. M. SOBOLÁ, Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates, Math. Comput. Simul., 55 (2001), pp. 271–280s.

[229] C. SOIZE AND R. G. GHANEM, Reduced chaos decomposition with random coefficients of vector-valued random variables and random fields, Computer Methods in Applied Mechanics and Engineering, 198 (2009), pp. 1926–1934.

[230] H. H. SONG, L. QIU, AND Y. ZHANG, NetQuest: A flexible framework for large-scale network measurement, Networking, IEEE/ACM Transactions on, 17 (2009), pp. 106–119.

[231] R. SRINIVASAN, Importance sampling - Applications in Communications and Detection, Springer-Verlag, New York, New York, USA, 2002.

[232] M. STEIN, Large sample properties of simulations using Latin hypercube sampling, Technometrics, 29 (1987), pp. 143–151.

[233] D. M. STEINBERG AND D. K. J. LIN, A construction method for orthogonal Latin hypercube designs, Biometrika, 93 (2006), pp. 279–288.

[234] D. E. STEPHENSON, E. M. HARTMAN, J. N. HARB, AND D. R. WHEELER, Modeling of particle-particle interactions in porous cathodes for lithium-ion batteries, J. Electrochem. Soc., 154 (2007), pp. A1146–A1155.

[235] W. SUBBER AND A. SARKAR, Domain decomposition method of stochastic pdes: a two-level scalable preconditioner, Journal of Physics: Conference Series, 341 (2012), p. 012033.

[236] V. R. SUBRAMANIAN, V. BOOVARAGAVAN, V. RAMADESIGAN, AND M. ARABANDI, Mathematical model reformulation for lithium-ion battery simulations: Galvanostatic boundary conditions, Journal of The Electrochemical Society, 156 (2009), pp. A260–A271.

[237] B. SUDRET, Global sensitivity analysis using polynomial chaos expansions, Reliability Engineering and System Safety, 93 (2008), pp. 964 – 979.

[238] P. L. TALLEC, Y.-H. D. ROECK, AND M. VIDRASCU, Domain-decomposition methods for large linearly elliptic three dimensional problems, J. Comp. Appl. Math., 34 (1991), pp. 93–117.

[239] S. TANG, C. YU, X. WANG, X. GUO, AND X. SI, Remaining useful life prediction of lithium-ion batteries based on the wiener process with measurement error, Energies, 7 (2014), p. 520.

[240] K. E. THOMAS, J. NEWMAN, AND R. M. DARLING, Mathematical modeling of Lithium batteries, in Advances in Lithium-Ion Batteries, W. Schalkwijk and B. Scrosati, eds., Springer USA, 2002, pp. 345–392.

[241] I. V. THORAT, Understanding performance-limiting mechanisms in lithium-ion batteries for high-rate applications, ProQuest, UMI Dissertation Publishing, 2011.

[242] I. V. THORAT, D. E. STEPHENSONA, N. A. ZACHARIASA, K. ZAGHIBB, J. N. HARBA, AND D. R. WHEELERA, Quantifying tortuosity in porous li-ion battery materials, Journal of Power Sources, 188 (2009), pp. 592–600.

[243] R. A. TODOR AND C. SCHWAB, Convergence rates for sparse chaos approximations of elliptic problems with stochastic coefficients, IMA Journal of Numerical Analysis, 27 (2007), pp. 232–261.

[244] A. TOSELLI AND A. KLAWONN, A feti domain decomposition method for maxwell's equations with discontinuous coefficients in two dimensions, tech. rep., Courant Institute, New York University, 1999.

[245] E. TYRTYSHNIKOV, Kronecker-product approximations for some function-related matrices, Linear Algebra and its Applications, 379 (2004), pp. 423–437.

[246] E. ULLMANN, Solution Strategies for Stochastic Finite Element Discretizations, PhD thesis, Technische Universitat Bergakademie Freiberg, 2008.

[247] L. O. VALOEN AND J. N. REIMERS, Transport properties of $lipf_6$-based li-ion battery electrolytes, Journal of The Electrochemical Society, 152 (2005), pp. A882–A891.

[248] X. WAN AND G. KARNIADAKIS, An adaptive multi-element generalized polynomial chaos method for stochastic differential equations, J. Comp. Phys., 209 (2005), pp. 617–642.

[249] C. Y. WANG, W. B. GU, AND B. Y. LIAW, Micromacroscopic coupled modeling of batteries and fuel cells: I. model development, J. Electrochem. Soc., 145 (1998), pp. 3407–3417.

[250] C. Y. WANG AND V. SRINIVASAN, Computational battery dynamics (cbd)electrochemical/thermal coupled modeling and multi-scale modeling, Journal of Power Sources, 110 (2002), pp. 364–376.

[251] Q. WANG, H. LI, X. HUANG, AND L. CHENZ, Determination of chemical diffusion coefficient of lithium ion in graphitized mesocarbon microbeads with potential relaxation technique, J. Electrochem. Soc., 148 (2001), pp. A737–A741.

[252] X. WANG AND I. H. SLOAN, Low discrepancy sequences in high dimensions: How well are their projections distributed?, Journal of Computational and Applied Mathematics, 213 (2008), pp. 366 – 386.

[253] R. WHEELER, AlgDesign. The R project for statistical computing. http://www.r-project.org/.

[254] N. WIENER, The homogenous chaos, American Journal of Mathematics, 60 (1938), pp. 897–936.

[255] P. WINKER AND K.-T. FANG, Monte Carlo and Quasi-Monte Carlo Methods 1996: Proceedings of a conference at the University of Salzburg, Austria, July 9–12, 1996, Springer New York, New York, NY, 1998, ch. Optimal U-Type Designs, pp. 436–448.

[256] H. P. WYNN, The sequential generation of D-optimum experimental designs, Ann. Math. Statist., 41 (1970), pp. 1655–1664.

[257] J. XIE, N. IMANISHI, T. MATSUMURA, A. HIRANO, Y. TAKEDA, AND O. YAMAMOTO, Orientation dependence of liion diffusion kinetics in licoo2 thin films prepared by rf magnetron sputtering, Solid State Ionics, 179 (2008), pp. 362–370.

[258] D. XIU, Numerical Methods for Stochastic Computations: A Spectral Method Approach, Princeton University Press, 2010.

[259] D. XIU AND J. HESTHAVEN, High-order collocation methods for differential equations with random inputs, SIAM J. Sci. Comput., 27 (2005), pp. 1118–1139.

[260] D. XIU AND G. KARNIADAKIS, The Wiener-Askey polynomial chaos for stochastic differential equations, SIAM Joural on Scientific Computing, 24 (2002), pp. 619–644.

[261] N. XUE, W. DU, A. GUPTA, W. SHYY, A. M. SASTRY, AND J. R. R. A. MARTINS, Optimization of a single lithium-ion battery cell with a gradient-based algorithm batteries and energy storage, J. Electrochem. Soc., 160 (2013), pp. A1071–A1078.

[262] L. YAN, L. GUO, AND D. XIU, Stochastic collocation algorithms using l_1-minimization, International Journal for Uncertainty Quantification, 2 (2012), pp. 279–293.

[263] J. YANG, B. YAN, J. YE, X. LI, Y. LIUA, AND H. YOUA, Carbon-coated licrtio4 electrode material promoting phase transition to reduce asymmetric polarization for lithium-ion batteries, Phys. Chem. Chem. Phys., 16 (2014), pp. 2882–2891.

[264] X. YANG AND G. E. KARNIADAKIS, Reweighted l_1 minimization method for stochastic elliptic differential equations, Journal of Computational Physics, 248 (2013), pp. 87 – 108.

[265] K. Q. YE, Orthogonal column Latin hypercubes and their application in computer experiments, Journal of the American Statistical Association, 93 (1998), pp. 1430–1439.

[266] K. Q. YE, W. LI, AND A. SUDJIANTO, Algorithmic construction of optimal symmetric Latin hypercube designs, Journal of Statistical Planning and Inference, 90 (2000), pp. 145 – 159.

[267] L. ZADEH, Fuzzy sets, Information and Control, 8 (1965), pp. 338 – 353.

[268] S. ZEIN, B. COLSON, AND F. GLINEUR, An efficient sampling method for regression-based polynomial chaos expansion, core discussion papers rp, Universit catholique de Louvain, Center for Operations Research and Econometrics (CORE), 2013.

[269] A. ZHANG, K.-T. FANG, R. LI, AND A. SUDJIANTO, Majorization framework for balanced lattice designs, Ann. Statist., 33 (2005), pp. 2837–2853.

[270] S. S. ZHANG, A review on the separators of liquid electrolyte li-ion batteries, Journal of Power Source, 164 (2007), pp. 351–364.

[271] T. ZHOU, A. NARAYAN, AND D. XIU, Weighted discrete least-squares polynomial approximation using randomized quadratures, Journal of Computational Physics, 298 (2015), pp. 787 – 800.

[272] T. ZHOU, A. NARAYAN, AND Z. XU, Multivariate discrete least-squares approximations with a new type of collocation grid, SIAM Journal on Scientific Computing, 36 (2014), pp. A2401–A2422.

[273] S. ZUGMANN, M. FLEISCHMANN, M. AMERELLER, R. GSCHWIND, H. WIEMHFER, AND H. GORES, Measurement of transference numbers for Lithium ion electrolytes via four different methods, a comparative study, Electrochimica Acta, 56 (2011), pp. 3926 – 3933.