

**Low-Rank, Multi-Fidelity Methods for Uncertainty  
Quantification of High-Dimensional Systems**

by

**H. R. Fairbanks**

B.A., Whitman College, 2010

M.S., University of Colorado Boulder, 2015

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 Applied Mathematics

2018

This thesis entitled:  
Low-Rank, Multi-Fidelity Methods for Uncertainty Quantification of High-Dimensional Systems  
written by H. R. Fairbanks  
has been approved for the Department of Applied Mathematics

---

Prof. Alireza Doostan

---

Prof. Gregory Beylkin

---

Prof. Stephen Becker

---

Prof. Jem Corcoran

---

Prof. Chris Ketelsen

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.

Fairbanks, H. R. (Ph.D., Applied Mathematics)

Low-Rank, Multi-Fidelity Methods for Uncertainty Quantification of High-Dimensional Systems

Thesis directed by Prof. Alireza Doostan

Characterizing and incorporating uncertainties when simulating physical phenomena is essential for improving model-based predictions. These uncertainties may stem from a lack of knowledge regarding the underlying physical processes or from imprecise measurements of quantities that describe properties of the physical system. Uncertainty quantification (UQ) is a tool that seeks to characterize the impact of these uncertainties on solutions of computational models, resulting in improved predictive models. In practice, these uncertainties are either treated as random parameters to inform the statistics of the solution of interest (forward UQ), or their statistics are inferred from noisy observations of the solutions (inverse UQ).

For systems exhibiting high-dimensional uncertainty, performing either forward or inverse UQ presents a significant computational challenge, as these methods require a large number forward solves of the high-fidelity model, that is, the model that accurately captures the physics of the problem. For large-scale problems, this may result in the need for a possibly infeasible number of simulations. Prominent methods have been developed to reduce the burdens related to these challenges, including multilevel Monte Carlo (for forward UQ) and low-rank approximations to the posterior covariance (for inverse UQ). However, these methods may still require many forward solves of the high-fidelity model.

To reduce the cost of performing UQ on high-dimensional systems, we apply multi-fidelity strategies to both the forward problem, in order to estimate moments of the quantity of interest, and inverse problem, to approximate the posterior covariance. In particular, we formulate multi-fidelity methods that exploit the low-rank structure of the solution of interest and utilize models of lower fidelity (which are computationally cheaper to simulate) than the intended high-fidelity model, in a nonintrusive manner. Doing so results in surrogate models that may have accuracies closer to

that of the high-fidelity model, yet have computational costs comparable to that of the low-fidelity models. Theoretical error analysis, cost comparisons, and numerical examples are provided to show the promise of these novel methods.

## Dedication

To my parents, for always encouraging me.

## Acknowledgements

There are many individuals who have been instrumental to the completion of this work. First and foremost, I would like to thank my advisor Alireza Doostan, whose hard work and dedication inspire me to be the best researcher I can be. Working with him has been an invaluable experience and I greatly appreciate the opportunity I have had to work with him. To my committee members: Stephen Becker, Greg Beylkin, Jem Corcoran, and Chris Ketelsen, I would like to thank you for your time and valuable input regarding this work. I would like to express my gratitude to the PSAAP II team at Stanford University, and, in particular, Gianluca Geraci, Lluís Jofre-Cruanyes, and Gianluca Iaccarino for their helpful input and efforts toward the UQ study presented in Chapter 3. I would like to thank Jerrad Hampton and Akil Narayan for their work regarding Chapter 2. This work would not be possible without the financial support of the United States Department of Energy under PSAAP II at Stanford University. I would like to thank Panayot Vassilevski and Sarah Osborn for advising me at Lawrence Livermore National Lab for the last two summers. There are many past and current graduate students at CU Boulder who I would like to recognize for their impact on my work. To Mohammad and Ji, for welcoming me into their research group. To my current office mates/trivia team Paul, Alec, Felix, and Michaela, for maintaining a fun and mostly productive work environment. To my close friends in the Applied Math department who have been pivotal during my time at CU: Meredith, Aly, Anna, Jeff, Ben, and Kerry; many years ago we started our work at CU, and it has been a pleasure sharing this journey with all of you. And finally I would like to thank my family. Although none have pursued studies in math, their continual encouragement and enthusiasm for my research has meant the world to me.

## Contents

### Chapter

<b>1</b>	<b>Introduction to Uncertainty Quantification of High-Dimensional Physical Systems</b>	<b>1</b>
1.1	Limitations and Approaches for High-Dimensional Uncertainty . . . . .	3
1.1.1	The Forward Problem . . . . .	3
1.1.2	The Inverse Problem . . . . .	5
1.2	Low-Rank, Multi-Fidelity Methods for Further Cost Reduction . . . . .	6
1.3	Contributions and Organization of this Thesis . . . . .	8
<b>2</b>	<b>Practical Error Bounds for a Non-intrusive Bi-fidelity Approach to Parametric/Stochastic Model Reduction</b>	<b>11</b>
2.1	Introduction . . . . .	11
2.1.1	Contribution of This Work . . . . .	14
2.2	Method Detail . . . . .	15
2.2.1	Low-Rank Factorization of Low-Fidelity Data . . . . .	17
2.2.2	High-Fidelity Approximation via Basis Update and Low-Fidelity Interpolation	18
2.2.3	Number of Low- and High-Fidelity Samples . . . . .	19
2.2.4	Convergence Analysis and Its Practical Application . . . . .	20
2.2.5	Estimating $\epsilon(\tau)$ Using Limited High-Fidelity Data . . . . .	23
2.3	Numerical Examples . . . . .	24
2.3.1	Test Case 1: Heat Driven Cavity Flow . . . . .	24

2.3.2	Test Case 2: Composite Beam . . . . .	34
2.4	Conclusions . . . . .	38
<b>3</b>	<b>A Bi-Fidelity Approximation for Uncertainty Quantification and Sensitivity Analysis of Ir-</b> <b>radiated Particle-Laden Turbulence</b>	<b>41</b>
3.1	Introduction . . . . .	41
3.1.1	Irradiated Particle-Laden Turbulent Flow . . . . .	41
3.1.2	Uncertainty Quantification for Complex, Large-Scale, High-Dimensional Sys- tems . . . . .	43
3.1.3	Objectives and Organization of the Work . . . . .	45
3.2	Physics Modeling and Numerical Method . . . . .	46
3.2.1	Variable-Density Turbulent Flow . . . . .	46
3.2.2	Lagrangian Particle Transport . . . . .	47
3.2.3	Radiative Heat Transfer . . . . .	49
3.2.4	Numerical Method . . . . .	49
3.3	Description of the Particle-Based Solar Receiver . . . . .	50
3.3.1	Computational Setup and Physical Parameters . . . . .	50
3.3.2	Uncertainties and Quantities of Interest . . . . .	50
3.3.3	Simulation Strategy and Ensemble Calculations . . . . .	53
3.3.4	Description of the High- and Low-Fidelity Models . . . . .	53
3.4	Bi-fidelity Approximation Strategy . . . . .	55
3.4.1	Formation of Bi-fidelity Approximation . . . . .	55
3.4.2	Theoretical Error Estimation of Bi-Fidelity Approximation . . . . .	57
3.4.3	Bi-Fidelity Approximation for Estimation of QoI Statistics . . . . .	59
3.5	Numerical Results of Bi-Fidelity Approximation . . . . .	60
3.5.1	QoI #1: Heat Flux Through the $\Delta x = 0.3L$ Plane . . . . .	61



3.5.2	QoI # 2: Spatially-Averaged and Point Estimates of $\Delta T/T_0$ Along Profile at Probe Location . . . . .	68
3.5.3	Computational Cost Comparisons of the Five Models . . . . .	79
3.6	Conclusions . . . . .	80
4	A Low-rank Control Variate for Multilevel Monte Carlo Simulation of High-dimensional Uncertain Systems . . . . .	<b>82</b>
4.1	Introduction . . . . .	82
4.1.1	Contributions of This Work . . . . .	85
4.2	Background on Multilevel Monte Carlo . . . . .	87
4.2.1	MSE of MLMC . . . . .	88
4.2.2	Number of Samples $N_\ell$ for $\hat{Y}_\ell$ . . . . .	88
4.2.3	MLMC Convergence Guarantee . . . . .	89
4.3	Multilevel Control Variates . . . . .	90
4.3.1	$Z_\ell$ Formulation . . . . .	90
4.3.2	$\mathbb{E}[Z_\ell]$ Estimation . . . . .	93
4.3.3	MSE Reduction via $W_\ell$ . . . . .	93
4.3.4	Number of Samples $\tilde{N}_\ell$ for $\hat{W}_\ell$ . . . . .	95
4.3.5	Number of Samples $N'_\ell$ for $\bar{Z}_\ell$ . . . . .	96
4.3.6	MLCV Implementation Details and Algorithms . . . . .	97
4.3.7	MLCV Cost Breakdown . . . . .	100
4.3.8	Discussion on Use of MLCV . . . . .	101
4.4	Numerical Results . . . . .	102
4.4.1	Test 1: Natural Frequency of a 2D Linear Elasticity Problem . . . . .	103
4.4.2	Test 2: Thermally Driven Cavity Flow . . . . .	109
4.4.3	Analysis . . . . .	114
4.5	Conclusion . . . . .	115

<b>5</b>	<b>A Bi-Fidelity Approximation for Linear Bayesian Inference</b>	<b>117</b>
5.1	Introduction . . . . .	117
5.1.1	Contributions of this work . . . . .	120
5.2	Bayesian Inference for Linear Gaussian Problems . . . . .	121
5.3	Low-Rank Approximation for Posterior Covariance Estimation . . . . .	123
5.4	Low-Rank, Bi-Fidelity Approximation for Posterior Covariance Estimation . . . . .	125
5.4.1	Theoretical Error Bound For Bi-Fidelity Approximation . . . . .	127
5.4.2	Comparison of Computational Cost . . . . .	132
5.5	Numerical Performance of Bi-Fidelity Approximation to Posterior Covariance . . . . .	134
5.5.1	Test Case # 1: $m_0 = 9$ . . . . .	136
5.5.2	Test Case # 2: $m_0 = 49$ . . . . .	145
5.6	Conclusions and Future Work . . . . .	153
	<b>Bibliography</b>	<b>154</b>
	<b>Appendix</b>	
<b>A</b>	<b>Matrix Interpolative Decomposition</b>	<b>164</b>
<b>B</b>	<b>Theoretical Framework for Bi-Fidelity Approximation Error Bound</b>	<b>166</b>
<b>C</b>	<b>Computational Methods for Computing Eigenvalue Decompositions</b>	<b>170</b>
C.1	Computing Eigenvalues and Eigenvectors . . . . .	170
C.1.1	Lanczos Algorithm . . . . .	170
C.1.2	Randomized Version of Lanczos Algorithm . . . . .	172
C.2	Comparison of Computational Cost . . . . .	174

## Tables

### Table

2.1	Description of the parameters of the composite cantilever beam model. The center of the holes are at $x = \{5, 15, 25, 35, 45\}$ . . . . .	35
3.1	Flow conditions at development section and physical properties. . . . .	51
3.2	List of random inputs with the corresponding ranges. All inputs are assumed to be uniformly distributed. . . . .	52
3.3	Statistics from Sparse PCE . . . . .	64
3.4	Statistics from sparse PCE . . . . .	74
4.1	Cost of MC, MLMC, and MLCV for Test 1 problem. The difference in cost of MC for $\varepsilon = 0.0005$ is due to the small number of samples used. . . . .	108
4.2	Cost of MC, MLMC, and MLCV for Test 2 problem. . . . .	114
5.1	Approximate Number of FLOPs for Different Low-Rank Approaches . . . . .	133
C.1	Number of FLOPs for Different Low-Rank Approaches . . . . .	175

## Figures

### Figure

2.1	Schematic of the heat driven cavity flow problem, reproduced from Figure 5 of [36]. .	26
2.2	Realizations of heat flux along the hot wall for a randomly selected input $\xi$ . Shown are the low-fidelity, high-fidelity, and rank $r = 10$ bi-fidelity estimates for the various low-fidelity models. . . . .	27
2.3	Histograms of low-fidelity and rank $r = 10$ bi-fidelity errors normalized by $\ \mathbf{H}\ $ for various low-fidelity models from 100 data samples. . . . .	28
2.4	Identification of error bound efficacy, i.e., the (average) ratio between the error estimated from (2.11) and the true error, for various low-fidelity meshes, approximation ranks $r$ , and sample sizes $n$ . For each pair $(n, r)$ the reported error ratio is the average of 30 ratios, each computed from an independent set of $n$ high- and low-fidelity samples via Algorithm 1. . . . .	30
2.5	$\hat{\epsilon}(\tau)$ from (2.16) as a function of $\tau$ for different low-fidelity model meshes with four combinations of bi-fidelity ranks $r = 5, 10$ and sample sizes $n = 10, 20$ . Values which optimize (2.11) for a rank $r$ bi-fidelity approximation are marked by black circles. Curves without optimal value markers displayed identify optimal bounds at $\hat{\epsilon}(\tau) = 0$ . . . . .	31
2.6	Minimum (over $k$ ) of error function $\rho_k(\tau)$ in (2.11b) for various low-fidelity models using an approximation rank $r = 10$ . In (a), minimizations occur (from coarsest to finest mesh) at $k = 13, 8, 9$ , and 85. In (b), minimizations occur (from coarsest to finest mesh) at $k = 7, 8, 9$ , and 9. Errors are normalized by $\ \mathbf{H}\ $ . . . . .	32

2.7	Minimum (over $k$ ) of error function $\rho_k(\tau)$ in (2.11b) for different low-fidelity models using different sample sizes $n$ . All bi-fidelity solutions are of rank $r = 10$ . In (a), minimizations occur (in ascending order of $n$ ) at $k = 13, 13, 7$ , and $7$ . In (b), minimizations occur (in ascending order of $n$ ) at $k = 85, 85, 9$ , and $9$ . Errors are normalized by $\ \mathbf{H}\ $ . . . . .	33
2.8	Estimate of (2.11) for various low-fidelity models and using different sample sizes $n$ . Errors are normalized by $\ \mathbf{H}\ $ . . . . .	33
2.9	Schematic of the cantilever beam (left) and its composite cross section (right). . . .	35
2.10	Finite element mesh used for high-fidelity simulation of the vertical displacement. . .	35
2.11	Realizations of vertical displacement for two randomly selected input parameters $\xi$ . Shown are the low-fidelity, high-fidelity, and rank $r = 1$ bi-fidelity estimates. . . . .	37
2.12	$\hat{\epsilon}(\tau)$ from (2.16) as a function of $\tau$ for different sample sizes $n$ . Values which optimize (2.11) for a rank $r = 1$ approximation are marked. . . . .	37
2.13	Histogram count for the relative error in the maximum displacement for low- and bi-fidelity models from 100 random realizations of the displacement. . . . .	38
2.14	Identification of error bound efficacy, i.e., the (average) ratio between the error estimated from (2.11) and the true error, for various sample sizes $n$ . For each $n$ the reported error ratio is the average of 30 ratios, each computed from an independent set of $n$ high- and low-fidelity samples via Algorithm 1. . . . .	39
2.15	Minimum (over $k$ ) of error function $\rho_k(\tau)$ in (2.11b) using different sample sizes $n$ . Minimization occurs at $k = 1$ for all four estimates. Errors are normalized by $\ \mathbf{H}\ $ . . .	39
3.1	Computational setup of the PSAAP II volumetric particle-based solar energy receiver. An isothermal periodic section (left domain) is utilized to generate fully developed particle-laden turbulent flow, which is used as inflow conditions for the second section (right domain) where the gas-particle mixture is irradiated perpendicularly to the flow direction from one the sides. . . . .	51

3.2	Detailed schematic of the radiated test section. The domain is $1.7L$ long and $W$ width, with the radiated region extending $\Delta x = L$ and $\Delta z = W$ starting at $\Delta x = 0.1L$ from the beginning of the section, and with a probe perpendicular to the flow and located $\Delta x = 0.3L$ downstream. . . . .	51
3.3	Turbulent gas phase Eulerian resolution of the fidelity levels designed; gas velocity as background. HF $540 \times 320 \times 320$ gridpoints (top), LF1 $108 \times 64 \times 64$ gridpoints (bottom left), LF2 $54 \times 32 \times 32$ gridpoints (bottom right). . . . .	54
3.4	Schematic for the formation of the BF approximation . . . . .	55
3.5	(a) Decay of normalized singular values of LF and HF matrices, using available data. (b) Error bound estimates for both BF approximations as a function of rank $r$ . For comparison, relative spectral error of the LF data are provided. . . . .	62
3.6	Error bound estimation in (3.19) using $\hat{\epsilon}(\tau)$ and rank $r = 6$ for varying values of $R$ , for (a) BF1 and (b) BF2 models. Values are based on 10 different sets of $R$ columns, where column selection not fully independent due to restricted number of HF samples. . . . .	63
3.7	Normalized total heat flux values for 17 independent simulations, from the five different models, where the BF approximation is of rank $r = 6$ . The BF approximations are more accurate with respect to the HF data than either LF QoI data. This data excludes simulations corresponding to basis data in the BF approximations. . . . .	64
3.8	Normalized histogram of the total normalized heat flux through the $x = \Delta 0.3L$ plane based on sparse PCE of (a) LF1 and BF1 QoI data, and (b) LF2 and BF2 QoI data. . . . .	66
3.9	Importance of input parameters calculated from sparse PCE coefficients on the (a) LF1, (b) BF1, (c) LF2, and (d) BF2 heat flux data. Starting with parameter $\xi_1$ at the top position, importance of each $\xi_i$ is provided in counterclockwise order with respect to increasing $i$ , with corresponding description provided in Table 3.2. Heat flux from the radiated wall ( $\xi_{12}$ ) and opposite wall ( $\xi_{13}$ ) to the fluid are the most important parameters as determined from all four model surrogates. . . . .	67

3.10	(a) Decay of normalized singular values of LF and HF matrices, using available $\Delta T/T_0$ data along the profile at the probe location. (b) Error bound estimates of the rank $r$ BF approximations for $\Delta T/T_0$ QoI, as a function of $r$ . For comparison, the LF1 and LF2 relative spectral errors are included as well. . . . .	69
3.11	Error bound estimation in (3.19) using $\hat{\epsilon}(\tau)$ and rank $r = 6$ for varying values of $R$ , for (a) BF1 and (b) BF2 models. Values are based on 10 different sets of $R$ columns, where column selection not fully independent due to restricted number of HF samples. . . . .	70
3.12	(a) Average $\Delta T/T_0$ profile calculated from available simulations for all five models, with rank $r = 6$ for the BF approximations. (b) $\ell_2$ error estimates of $\Delta T/T_0$ at each point along the profile. . . . .	71
3.13	$\Delta T/T_0$ values of 17 independent simulations for the QoI of (a) spatial mean along profile at probe location, and points (b) $y/W = 0.5$ , (c) $y/W = 0.1$ , and (d) $y/W = 0.05$ , along probe profile. Simulated values are from the five different models, where the BF approximation is of rank $r = 6$ . . . . .	72
3.14	Normalized histograms of the LF1 and BF1 surrogate models (left column) and LF2 and BF2 surrogate models (right column) for the four $\Delta T/T_0$ QoIs: mean ((a) and (b)), $y/W = 0.5$ ((c) and (d)), $y/W = 0.1$ ((e) and (f)), and $y/W = 0.05$ ((g) and (h)). Histograms formed from 25,000 samples of the sparse PCE surrogates. . . . .	75
3.15	Importance of input parameters for the spatial mean $\Delta T/T_0$ QoI from sparse PCE coefficients of the (a) LF1 (b) BF1, (c) LF2, and (d) BF2 surrogate models. Starting with parameter $\xi_1$ at the top position, importance of each $\xi_i$ is provided in counter-clockwise order with respect to increasing $i$ , with corresponding description provided in Table 3.2. . . . .	77

3.16	Importance of input parameters from sparse PCE coefficients from the BF2 model for the $\Delta T/T_0$ QoIs of (a) spatial mean along profile at probe location (b) $y/W = 0.5$ along profile at probe location, (c) $y/W = 0.1$ along profile at probe location, and (d) $y/W = 0.05$ along profile at probe location. Starting with parameter $\xi_1$ at the top position, importance of each $\xi_i$ is provided in counterclockwise order with respect to increasing $i$ , with corresponding description provided in Table 3.2. . . . .	78
3.17	Number of core-hours on Mira (ALCF) [2] to obtain $N$ simulated values from each model. Markers provide number of core-hours for $N = 128$ and $N = 256$ , to provide the cost of the simulated values generated for this study. . . . .	80
4.1	(a) Schematic of the L-shaped domain. (b)-(f) Meshes on the five different levels with increasing resolution. . . . .	105
4.2	Change of the correlation $\rho_\ell^2$ in (4.22) as a function of the ID rank $r$ for (left) case $A$ and (right) case $B$ . . . . .	105
4.3	(Left) Convergence of the sample mean of $Y_\ell$ and $W_\ell$ , while that of $Q_\ell$ remains relatively constant. (Right) Convergence of the sample variance of $Y_\ell$ in comparison to that of $Q_\ell$ . Superscripts $A$ and $B$ refer to the case $A$ and $B$ of the L-shaped domain problem, respectively. . . . .	106
4.4	(Left) MSE reduction factor (MSERF) between $\hat{W}_\ell$ and $\hat{Y}_\ell$ , as well as the value of $\rho_\ell^2$ on each level. (Right) Total number of samples needed to achieve the sampling error $\varepsilon = 0.0005$ for MLMC and MLCV. MLCV sample size includes the number required to determine $\bar{Z}_\ell$ . Superscripts $A$ and $B$ refer to the case $A$ and $B$ of the L-shaped domain problem, respectively. . . . .	107
4.5	(Left) MSE for $\hat{Y}_\ell$ and $\hat{W}_\ell$ . (Right) Comparison of the relative errors (4.31) and (4.32) in estimating the mean QoI at each level for MLMC and MLCV. Superscripts $A$ and $B$ refer to the case $A$ and $B$ of the L-shaped domain problem, respectively. . . . .	109



4.6	Schematic of Test 2, the thermally driven flow problem with random temperature along the hot (west) wall and spatially varying stochastic temperature along the cold (east) wall. . . . .	110
4.7	Change of the correlation $\rho_\ell^2$ in (4.22) as a function of the ID rank $r$ for Test 2. . .	111
4.8	(Left) Convergence of the sample mean of $Y_\ell$ and $W_\ell$ , while that of $Q_\ell$ remains relatively constant. (Right) Convergence of the sample variance of $Y_\ell$ in comparison to that of $Q_\ell$ . . . . .	112
4.9	(Left) MSE reduction factor (MSERF) between $\hat{W}_\ell$ and $\hat{Y}_\ell$ , as well as the value of $\rho_\ell^2$ on each level. (Right) Total number of samples needed to attain the sampling error for MLMC and MLCV when $\varepsilon = 0.001$ . MLCV sample size includes the number required to determine $\bar{Z}_\ell$ . . . . .	113
4.10	(Left) MSE for $\hat{Y}_\ell$ and $\hat{W}_\ell$ . (Right) Comparison of the relative errors (4.31) and (4.32) in estimating the mean of QoI at each level for MLMC and MLCV. . . . .	114
5.1	(a) Diffusion field over the physical domain $D$ . (b) True initial value of $\mathbf{u} = \mathbf{u}_0$ over $D$ at time $t = 0$ . . . . .	135
5.2	Solution $\mathbf{u}$ at final time $t = T$ , when using a reference mesh grid of size $128 \times 128$ . .	137
5.3	Numerical solutions (excluding boundary) to the PDE in (5.25) for (a) the fine grid mesh and (b) the coarse grid mesh. Red “x” markers indicate where observations are extracted. . . . .	138
5.4	L-curve comparing the error term and regularization term of (5.8). Optimal point of $\alpha = 0.07$ selected as balances the two terms. . . . .	138
5.5	Decay of the eigenvalues of the data misfit of the Hessian for regularization parameter values (a) $\alpha = 0.07$ and (b) $\alpha = 500$ . High-fidelity refers the eigenvalues $\mathbf{\Lambda}_r$ , as in (5.10), and low-fidelity refers to eigenvalues $\tilde{\mathbf{\Lambda}}_r$ , as in (5.14). . . . .	139

5.6	Relative error of the low-rank and low-rank, bi-fidelity approximations as in (5.28) for the scenarios (a) $\alpha = 0.07$ and (b) $\alpha = 500$ . Plots on the left correspond to the use of direct SVD to calculate the eigenvalue decompositions, and plots on the right correspond to using randomized approaches. . . . .	141
5.7	Euclidean distance between the true maximum point coordinate (0.67,0.67) and the maximum point coordinate as calculated from the posterior mean (via MAP estimation) in (5.8). Three methods are compared: low-rank (high-fidelity), new bi-fidelity, and low-rank, low-fidelity approaches, as a function of rank $r$ . . . . .	143
5.8	Cost comparisons between the low-rank (high-fidelity) method and our bi-fidelity approximation, as a function of rank $r$ , for (a), (c) randomized SVD and (b), (d) randomized Lanczos for eigenvalue decomposition calculations. These estimates are based on FLOP decompositions. (a)-(b) compare the cost when using only $n_t = 1$ time steps and (c)-(d) when using $n_t = 45$ time steps. Dashed lines indicate the cost difference of the two methods. . . . .	144
5.9	Numerical solutions (excluding boundary) to the PDE in (5.25) for (a) the fine grid mesh and (b) the coarse grid mesh. Red “x” markers indicate where observations are extracted. . . . .	146
5.10	L-curve comparing the error term and regularization term of (5.8). Optimal point of $\alpha = 0.014$ selected as balances the two terms. . . . .	146
5.11	Decay of the eigenvalues of the data misfit of the Hessian for regularization parameter values (a) $\alpha = 0.014$ and (b) $\alpha = 200$ . High-fidelity refers the eigenvalues $\mathbf{\Lambda}_r$ , as in (5.10), and low-fidelity refers to eigenvalues $\tilde{\mathbf{\Lambda}}_r$ , as in (5.14). . . . .	147
5.12	Relative error of the low-rank and low-rank, bi-fidelity approximations as in (5.28) for the scenarios (a) $\alpha = 0.014$ and (b) $\alpha = 200$ . Plots on the left correspond to the use of direct SVD to calculate the eigenvalue decompositions, and plots on the right correspond to using randomized approaches. . . . .	148

5.13	Euclidean distance between the true maximum point coordinate (0.67,0.67) and the maximum point coordinate as calculated from the posterior mean (via MAP estimation) in (5.8). Three methods are compared: low-rank (high-fidelity), new bi-fidelity, and low-rank, low-fidelity approaches, as a function of rank $r$ .	151
5.14	Cost comparisons between the low-rank (high-fidelity) method and our bi-fidelity approximation, as a function of rank $r$ , for (a), (c) randomized SVD and (b), (d) randomized Lanczos for eigenvalue decomposition calculations. These estimates are based on FLOP decompositions. (a)-(b) compare the cost when using only $n_t = 1$ time steps and (c)-(d) when using $n_t = 45$ time steps. Dashed lines indicate the cost difference of the two methods.	152

## Chapter 1

### Introduction to Uncertainty Quantification of High-Dimensional Physical Systems

When modeling physical phenomena, there are typically numerous uncertainties present. These uncertainties may stem from a lack of knowledge regarding the underlying physical processes or from imprecise measurements of quantities that describe properties of the physical system. Uncertainty quantification (UQ) is a tool that seeks to characterize the impact of these uncertainties on solutions of computational models, resulting in improved predictive models. In particular, by treating these uncertainties as randomly varying parameters, instead of fixed values as in a deterministic setting, resulting solutions of these systems are random, and UQ methods are tasked with calculating the statistics of these solutions. With the introduction of these uncertain input parameters, standard deterministic models must either be altered to incorporate them, referred to as intrusive methods, or, alternatively, require repeated simulations of the deterministic model with different values of the input parameter according to its probability distribution, referred to as non-intrusive methods.

In this work we consider two types of UQ: the forward problem and the inverse problem. The forward problem aims to determine the statistics of the the output of the model, which we refer to as the quantity of interest (QoI). For the forward problem, the probability distribution of the uncertain parameters is assumed to be known. The inverse problem, on the other hand, aims to improve how we define the uncertain parameters, given observations of the QoI. An approach for the inverse problem is to use Bayesian inference, where the uncertain parameters, or as they are

termed for statistical inference, unknown parameters, are characterized with an assumed probability distribution that is then updated using observational data.

A central application of this work is the Predictive Science Academic Alliance Program (PSAAP) II, where the focus is developing numerical methods to model particle-based solar receivers [1]. This type of concentrated solar power system relies on heat transport between fluid and particles as a means to harness solar energy. While highly efficient, the underlying physical processes are not well understood. The aim of the program is to effectively model the system to improve the energy efficiency by optimizing power output. Outside the scope of UQ, forming models of such complex systems is no simple task. This multi-physics problem requires incorporating particle tracking for  $\mathcal{O}(10^6)$  particles, turbulent flows, and an external radiation source to mimic the physical processes inherent in a particle-based solar receiver. The coupling of these processes results in an extremely computationally demanding numerical model to form predictions for this type of concentrated solar power system. For example, a typical converged simulation for a small-scale particle-based solar receiver requires about 500 thousand core-hours.

In addition to having a complex system, there are many uncertainties within this physical process, including the particle properties and radiation intensity. Altogether, this could amount to  $\mathcal{O}(10^6)$  uncertainties. Accounting for and characterizing the effect of these uncertainties is of utmost importance for improved predictive methods as well as identifying the QoI's sensitivity to parameter variations. However, as modeling this process is already computationally demanding, i.e., it requires significant computational resources to complete a single simulation, the addition of a large number of uncertainties makes performing UQ a difficult task. Due to this, the main focus of this work is developing methods to alleviate the high computational demands of performing UQ in the presence of high-dimensional uncertainty. We remark here that the concept of *high-dimension* is problem dependent, but is generally considered to be the dimension  $d$  at which the required number of simulations is infeasible for UQ methods, where the cost of performing UQ scales poorly as a function of  $d$ , e.g., exponentially.

As a primary motivation for this work, PSAAP II requires non-intrusive UQ methods. As

such, and due to the advantageous nature of utilizing non-intrusive methods (e.g., they allow for legacy codes), the focus of this work is placed on the formation of non-intrusive UQ algorithms for high-dimensional uncertainty.

## 1.1 Limitations and Approaches for High-Dimensional Uncertainty

Non-intrusive UQ methods require performing repeated simulations of the model of interest. For problems exhibiting high-dimensional uncertainty many strategies for performing UQ are problematic, as they suffer from the curse of dimensionality, where the number of simulations required to meet a desired error tolerance grows rapidly as a function of the number of uncertainties. For computationally demanding models, such as in the PSAAP II application, a large number of model simulations is infeasible, restricting the types of methods that can be applied to perform UQ. The curse of dimensionality affects both the forward and inverse problems in different ways. In the following two subsections, we address the specific issues that arise for forward and inverse UQ with brief discussions of leading methodologies for addressing this issue.

### 1.1.1 The Forward Problem

The goal of the forward problem is estimating the statistics of the QoI; particularly, we are interested in the mean, variance, and PDF of the QoI, dependent on the statistics of the random  $d$ -dimensional uncertain parameter  $\boldsymbol{\xi} = (\xi_1, \dots, \xi_d)$ . Each component  $\xi_i$  represents a different uncertainty in the system, e.g.,  $\xi_1$  is the parameter that describes the radiation intensity as in the PSAAP II problem, and the joint PDF  $\rho(\boldsymbol{\xi})$  is assumed to be either known, or be well approximated.

Several non-intrusive methods, e.g., stochastic collocation [137, 88] and polynomial chaos expansions (PCEs) [48, 136, 29], have been developed and proven successful in various UQ applications. However, the computational cost of these methods grows rapidly as a function of the number of random variables  $d$ . For example, stochastic collocation is a numerical integration technique for estimating the mean value of the QoI, where the number of required simulations to estimate the integral grows exponentially in  $d$ . Thus these approaches are not feasible for large  $d$ . Extensions

of these works, such as sparse grid stochastic collocation [135] and sparse approximations to PCEs [29, 60] greatly alleviate the curse of dimensionality observed for stochastic collocation and PCEs, respectively; however, the size of  $d$  will still affect the required number of simulations.

An alternative class of techniques that does not suffer from the curse of dimensionality is the Monte Carlo (MC) method and its variants. For these methods, the statistics of the QoI are estimated using an ensemble of random realizations of the QoI [58]. The cost of such estimations, while maybe prohibitive, is *formally* independent of the number of input variables.

Let  $Q = Q(\boldsymbol{\xi})$  denote a scalar-valued QoI depending on  $\boldsymbol{\xi}$  and let  $Q_L$  be the computational approximation of  $Q$  with resolution parameter  $L$ , e.g., related to grid size. We let  $M_L$  denote the number of deterministic degrees of freedom, e.g., number of grid points in a finite element model, controlling the accuracy of  $Q_L$  relative to  $Q$ . The goal is to approximate the statistics of  $Q$ , e.g., the mean of  $Q$ ,  $\mathbb{E}[Q]$ , using the realizations of  $Q_L$ . Given a set of  $N$  input samples drawn randomly according to the joint PDF, each denoted by  $\boldsymbol{\xi}^{(i)}$ , and the corresponding realizations of the QoI, given by  $Q_L^{(i)} = Q(\boldsymbol{\xi}^{(i)})$ , the MC approximation of  $\mathbb{E}[Q]$  is defined as

$$\hat{Q}_L := \frac{1}{N} \sum_{i=1}^N Q_L^{(i)}. \quad (1.1)$$

It can be shown that the mean square error (MSE) of this approximation, defined as  $\mathbb{E}[(\hat{Q}_L - Q)^2]$ , may be decomposed as

$$\mathbb{E}[(\hat{Q}_L - Q)^2] = \frac{\mathbb{V}[Q_L]}{N} + (\mathbb{E}[Q_L - Q])^2.$$

The sampling error  $\mathbb{V}[Q_L]/N$  is controlled by the variance of  $Q_L$  and the number of samples, and the discretization error  $(\mathbb{E}[Q_L - Q])^2$  measures how closely the model simulates the true solution. As can be seen from this decomposition, the sampling error decays slowly as a function of  $N$ , but with a rate that is independent of the dimension  $d$ , implying that the standard MC simulation does not formally suffer from the curse of dimensionality.

Despite the fact that MC does not suffer from the curse of dimensionality, it is often too cost prohibitive when integrating high-dimensional uncertainty with high-cost simulations. This leads us to consider two options to improve the cost of MC: (i) formulating a model reduction technique

to approximate the QoI with reduced simulation cost, and (ii) employing a variance reduction technique. Variance reduction techniques, such as control variates or importance sampling [58, 8], lower the sampling error, thus requiring fewer samples of  $Q_L$  to meet a desired MSE tolerance. Often, both (i) and (ii) are achieved in unison, i.e., researchers develop a reduced order model and combine it with a variance reduction technique to obtain a desired MSE with improved computation time [124, 131, 93, 99].

An example of a control variate that has been widely recognized for its promising results is multilevel Monte Carlo (MLMC) [65, 49, 24, 128], which constructs a sequence of control variates based on approximations of  $Q$  on a set of models that are cheaper to simulate than the one for  $Q_L$ , hence the term *multilevel*. A common example of a cheaper model is to approximate  $Q$  on coarser grids with number of degrees of freedom smaller than  $M_L$ . Other work in variance reduction techniques that rely on a multilevel scheme can be found in [131], where the sequence of control variates is constructed based on the use of a reduced basis in the underlying finite element model; and in [93], where the sequence of control variates, determined by solving the partial differential equation (PDE) with smooth realizations of the random input field, is applied to reduce the variance of MLMC.

### 1.1.2 The Inverse Problem

A common way to approach the inverse problem is by using Bayesian inference. In a computational setting the goal is to use observational data of the QoI to infer the statistics of the  $N$ -dimensional unknown parameter  $\mathbf{x}$ . Initially an approximate probability distribution is given to the unknown parameter, referred to as the *prior distribution*, and then noisy observational data is used to update this information via Bayes' theorem, such that the probability distribution of the input parameter is more accurately informed. The updated probability distribution is referred to as the *posterior distribution*. For scenarios of interest to this work for the inverse problem, the forward map is linear, and both the observational noise and prior distribution are Gaussian, thus the posterior distribution will be Gaussian as well. Accordingly, the problem is simplified to estimating



the posterior mean and covariance, which is done via maximum *a posteriori* (MAP) estimation.

Of significant challenge is calculating the posterior covariance as it requires forming and inverting an  $N \times N$  matrix. In this work, the number of unknowns  $N$  corresponds to the degrees of freedom in the numerical discretization of the governing equations. For large-scale problems, where the discretization is finely resolved,  $N$  is too large for direct inversion. In particular, the computational complexity for the matrix inversion is  $\mathcal{O}(N^3)$ . Recent works in [40, 18, 19] reformulate this matrix inversion problem into a low-rank eigenvalue decomposition problem. By doing so, they are able to reduce the computational cost of estimating the posterior covariance. Without going into detail, this is done by reducing the total number of simulations required of the model.

## 1.2 Low-Rank, Multi-Fidelity Methods for Further Cost Reduction

Significant progress for cost reduction has been made with multilevel and low-rank methods as discussed in Section 1.1. Both methods achieve cost reduction by reducing the number of necessary simulations of the high-fidelity model. The high-fidelity model refers to the solver which accurately simulates the physical system. Typically these models are expensive to run, and performing many simulations may be infeasible. As an example, we may use a finite element solver with a high-resolution for the spatial discretization scheme (fine grid solver). In contrast, low-fidelity models are less accurate, yet cheaper to simulate. Examples of such models include reduced resolution of the spatial discretization scheme (coarse grid solver), reduced number of time steps for the convergence of steady-state solutions, or models with simplified physics as compared to the high-fidelity model.

In recent years, there has been growing interest in multi-fidelity methods for performing UQ, that is, methods that leverage models of lower fidelity than intended high-fidelity model. In some cases, the combination low-fidelity and high-fidelity model solutions results in models with accuracies similar to that of the high-fidelity solution and computational cost closer to that of the low-fidelity model. Past work in multi-fidelity methods has been done in other fields, such as optimization, to reduce the cost of performing a large number of high-fidelity model simulations. In optimization, multi-fidelity methods utilize a low-fidelity model to search the parameter space

over which the optimization is occurring, with an aim of reducing the size of this search space when optimizing with respect to the high-fidelity model [87, 45]. In several cases, non-intrusive multi-fidelity methods are constructed by estimating the QoI from a low-fidelity model, and applying an additive or multiplicative correction term that is calculated using a few corresponding low-fidelity and high-fidelity simulations [26]. Examples of this include multi-fidelity PCEs [32, 92], MLMC (see Section 1.1.1), and multi-level compressed sensing [14]. From the geosciences, kriging and co-kriging have been popular methods to form a surrogate to the high-fidelity model via Gaussian process regression between low-fidelity model simulations and high-fidelity model simulations [72, 107]. For a review of multi-fidelity methods see, e.g., [99, 39].

In this work, we focus our attention on a specific class of multi-fidelity methods, particularly methods that exploit low-rank structure of the solution space, and take advantage of the similarities between the spaces described by the low-fidelity and high-fidelity models. The underlying assumption is that the high-dimensional space of high-fidelity solutions exist on a lower-dimensional manifold, and may thus be well approximated with fewer simulations of the high-fidelity model, resulting in further cost reduction. In many applications, the aim of dimension reduction is to identify the underlying data structure, resulting in improved efficiency for storage and data analysis. This can be done using techniques such as singular value decompositions, QR factorizations, or interpolative decompositions; we note, for high-dimensional datasets, randomized versions of these have become increasingly popular due to their improved efficiency [57, 84, 83]. Of interest in this work, is using these dimension reduction techniques to identify a reduced basis with which we may accurately approximate high-fidelity simulations.

Much of the work in this class of low-rank approximation methods corresponds to techniques referred to as reduced basis methods [94, 104, 106, 114]. These methods rely on first identifying a small set of spatial/temporal basis vectors of the solution from many high-fidelity model evaluations, where the reduced basis may be formed via methods such as proper orthogonal decomposition or singular value decomposition. Second, the solution to the high-fidelity model is projected onto the identified reduced basis, instead of, for instance, standard finite element basis. The first step

typically requires an expensive offline stage, which may be computationally infeasible for large-scale problems. The second step often requires modifications to the existing PDE solvers and thus cannot be conveniently implemented. Unlike these types of reduced basis methods, our work relies on non-intrusive, low-cost formulations that do not rely on an expensive offline stage to identify a reduced basis. Rather, we aim to exploit low-fidelity models to identify a reduced basis scheme for the high-fidelity model, resulting in significant cost reduction while maintaining reasonable accuracy.

### 1.3 Contributions and Organization of this Thesis

This thesis is organized into five chapters, where Chapters 2-5 are presented as independent works, with independently defined notation. Chapters 2-4 address our low-rank, multi-fidelity methods for the cost reduction for the forward problem when the uncertainty is high-dimensional. In particular, the work of these three chapters centers on a bi-fidelity approximation, which relies on PDE solutions from a high-fidelity and low-fidelity model, to obtain a low-rank, low-cost approximation of the high-fidelity solution.

The construction of this bi-fidelity approximation, inspired by the work in [91, 142, 28, 27], consists of four main steps. In the first step,  $N$  realizations of the low-fidelity solution are obtained. In the second step, we identify a reduced basis, say of size  $r \ll M_L$ , for the low-fidelity solution, together with an interpolation rule that gives an arbitrary realization of the low-fidelity solution in that basis. For this purpose, we borrow ideas from matrix interpolative decomposition as presented in [23]. The third step entails the identification of a reduced basis for the high-fidelity solution. For this, we generate the high-fidelity counterpart of the low-fidelity reduced basis, which requires  $r$  realizations of the high-fidelity solution. In the fourth step, we apply the same interpolation rule, as for the low-fidelity solution, to generate the low-rank approximation of its high-fidelity counterpart (and subsequently  $Q_L$ ). This results in a reduced model formulation of the high-fidelity solutions.

In practice this bi-fidelity approximation has been shown to be a favorable surrogate for the high-fidelity model when using MC sampling [27]. However, a significant challenge is determining, *a priori*, how the error of this approximation scales with respect to the high-fidelity solution.

Chapter 2 develops a theoretical framework with which we may bound the error of this bi-fidelity approximation with respect to data derived with the high-fidelity model. From this, we may estimate the error of this bi-fidelity approximation while requiring access to many low-fidelity solutions and a few high-fidelity solutions, resulting in minimal added cost. Numerical testing of this error estimate is presented to highlight its applicability. This is work in collaboration with Dr. J. Hampton (University of Colorado Boulder), Prof. A. Doostan (University of Colorado Boulder), and Prof. A. Narayan (University of Utah) [62].

In the case where the bi-fidelity error is small, this approximation may be used as a surrogate to simulations of the high-fidelity model to estimate the statistics of the QoI with similar accuracies to the high-fidelity simulations, but with significantly reduced cost. Chapter 3 explores the use of this bi-fidelity approximation combined with PCEs, as a surrogate for the PSAAP II application. We consider two surrogates in this work formed between a high-fidelity model and two different low-fidelity models. With these surrogates, we are able to more accurately estimate the statistics of multiple thermal QoIs than the corresponding low-fidelity models. The work of this chapter is in collaboration with Dr. L. Jofre (Stanford University), Dr. G. Geraci (Sandia National Laboratory), Prof. A. Doostan (University of Colorado Boulder), and Prof. G. Iaccarino (Stanford University); in particular, the content describing the PSAAP II model is provided by Dr. L. Jofre [37].

In the case that this bi-fidelity approximation may not serve as a favorable estimate for direct use, it can be used as a control variate on MC-based methods. In Chapter 4 we extend the use of this bi-fidelity approximation to serve as a control variate for MLMC, resulting in further cost reduction. Through the use of this error estimate and variance reduction technique, it can be shown that for applications exhibiting a low rank, this multi-fidelity technique promotes fast propagation of high-dimensional uncertainties. This is work in collaboration with Prof. C. Ketelsen (University of Colorado Boulder), Prof. A. Doostan (University of Colorado Boulder), and Prof. G. Iaccarino (Stanford University) [36].

Chapter 5 investigates the use of a new low-rank, bi-fidelity approximation for the inverse problem for further cost reduction. In this work we exploit low-rank structure and similarities

in the high-fidelity and low-fidelity solvers. Inspired by the work in [40, 18] we consider a similar formulation for estimating the posterior covariance, where instead of direct inversion of an  $N \times N$  matrix, we seek a low-rank eigenvalue decomposition. However, unlike the work in [40, 18] we consider a bi-fidelity approach. In particular, we use a low-fidelity model to calculate the low-rank eigenvalue decomposition, which is cheaper in comparison to using the high-fidelity model. Then, employing ideas from the Nyström approximation [134, 42], we approximate the high-fidelity eigenvectors using the eigenvectors and eigenvalues derived from the low-fidelity model. From theoretical and numerical results, we show that this bi-fidelity approximation is similar in accuracy to the high-fidelity approach for low-rank systems, but cheaper to construct.

## Chapter 2

### Practical Error Bounds for a Non-intrusive Bi-fidelity Approach to Parametric/Stochastic Model Reduction

#### 2.1 Introduction

In many practical contexts, an ideal computational method accurately recovers the physical phenomena that it is tasked to model, and does so in a computationally efficient manner that allows repeated calculations of quantities of interest (QoI's) for a large number of different input parameters. When this is possible, the QoI's may be understood across a range of input parameters, or similarly for certain distributions of input parameters. The latter case is connected with the field of parametric uncertainty quantification (UQ), and is fundamental for understanding the propagation of uncertainty through complex models often via the construction of the map between the inputs and QoI's. A growing area of interest in model-based simulations is a fast construction of such a mapping, which also forms our motivation here. For physical systems that are expensive to simulate, such a task may not be straightforward and may provide a significant challenge as it often requires a tradeoff between computational cost and accuracy.

In practice, the model which accurately represents the underlying physics – the *high-fidelity* model – may be expensive, and many repeated evaluations are likely computationally infeasible. Alternatively, one may consider the use of a *low-fidelity* model: This is a model with lower accuracy which is cheaper to evaluate – relative to the high-fidelity model – but not necessarily as accurate. The topic of this paper is a *bi-fidelity* approach that leverages both models and inherits the strengths of each.

In this work, we assume the existence of a low- and a high-fidelity model that each identify the same QoI in a discrete sense, i.e., as a vector, possibly with different sizes and levels of accuracy. For example, a fine spatial discretization of a heat equation, while costly, may lead to accurate temperature predictions, but a coarse spatial discretization gives inaccurate estimates that are generated more quickly. The notion of accuracy here is contextual, both with respect to the metric and the desired quantitative tolerance. Mathematically, we take the high-fidelity model as one that gives an approximation to the *truth* to within a desired accuracy and ideally with smallest possible computational cost. It is critical that we assume such a high-fidelity model is still expensive to evaluate, so that methods that would require many simulations from this model are impractical. We also assume that it is (much) cheaper to sample from the low-fidelity model so that compiling a larger number of samples from the low-fidelity model is possible. The goal of a (non-intrusive) bi-fidelity, or more generally multi-fidelity, model is to use relatively few samples from the high-fidelity model and a larger number of samples from the low-fidelity model to generate approximations to the QoI – at arbitrary input samples – with accuracies comparable to that of the high-fidelity model.

Early work in multi-fidelity methods for parametric problems has a decidedly statistical flavor. Such work emerged in the field of geosciences with the application of Gaussian process regression – also known as kriging or co-kriging in the multivariate case – in a multi-fidelity setting. In [72], an autoregressive scheme is considered to generate a Gaussian process approximation of the output of the most expensive model from nested observations of multiple, less expensive models. In this autoregressive scheme, a given model is approximated by the lower-fidelity model through a multiplicative constant and an additive Gaussian process correction term. In a related work, [107] creates a Gaussian process surrogate to the low-fidelity model, which is subsequently adjusted via samples of the high-fidelity model. The adjustment is done by training a similar autoregressive model as in [72], consisting in a linear – in input parameters – mapping of the low-fidelity surrogate and an additive discrepancy term modeled as a Gaussian process. The resulting high-fidelity surrogate is also a Gaussian process. Extensions of multi-fidelity Gaussian process modeling for

applications in optimization are explored in [41, 76], and recursive methods for improved accuracy or computational complexity are developed in [73, 53, 52, 102, 101, 97], among other work.

There has also been significant work over the last several years on reformulating other prominent UQ methods in a multi-fidelity (or multi-level) framework. These methods generally rely on a relatively large number of low-fidelity samples along with a (much) smaller number of high-fidelity samples to build an additive and/or multiplicative correction of the low-fidelity model. The adjusted low-fidelity model will then serve as a multi-fidelity approximation to the high-fidelity model. In particular, in multi-fidelity polynomial chaos expansions (PCEs) [32, 92, 96, 95], the high-fidelity solution is estimated by first determining the PCE of the low-fidelity solution, and subsequently the PCE of a correction term is generated from a small set of low-fidelity and high-fidelity samples. In multi-level Monte Carlo (MC) [65, 49, 24], a sequence of fine-to-coarse level (or grid) models are considered and the expectation of the QoI is estimated via a telescoping summation. These methods apply MC to the coarsest level QoI, and then add the MC estimates of the QoI differences, a.k.a. corrections, over each pair of consecutive grids. Assuming the corrections over finer grids become smaller, fewer realizations of the finer (high-fidelity) grid QoI are needed to compute the correction expectations, thus resulting in a reduced overall cost.

With an aim of constructing a reduced polynomial representation of the high-fidelity solution, the work in [28, 47] builds a small size polynomial basis – in the random inputs – identified by a Karhunen-Loève expansion of the low-fidelity solution. The high-fidelity solution is then approximated via Galerkin projection of the governing equations on the span of the reduced polynomial basis – as opposed to the standard PCE basis – resulting in a much smaller size system of equations to solve. The key assumption in this bi-fidelity construction is that the largest eigenvalues of the covariance matrix of the discrete solution decay quickly; that is, the covariance matrix and, therefore, the solution are low-rank. A non-intrusive implementation of this work is found in [74].

Of most relevance to the approach of this study, [91, 142] considers the low-rank approxima-



tion of the discretized QoI  $\mathbf{v} \in \mathbb{R}^M$ ,

$$\mathbf{v}(\boldsymbol{\xi}) \approx \sum_{\ell=1}^r \mathbf{v}(\boldsymbol{\xi}^{(\ell)}) c_{\ell}(\boldsymbol{\xi}). \quad (2.1)$$

The approximation above is of the same type constructed in so-called *reduced basis* representations in the (parametric) model reduction literature [5, 94, 104, 106, 114]. However, the procedure considered in [91, 142] and here is disparate from the algorithmic strategies in the reduced basis community. In (2.1),  $\mathbf{v}$  at an arbitrary parametric input  $\boldsymbol{\xi} \in \mathbb{R}^d$  is represented in a basis consisting of  $r \ll M$  realizations of  $\mathbf{v}$  – corresponding to selected input samples  $\boldsymbol{\xi}^{(\ell)}$  – via the coefficients  $c_{\ell}(\boldsymbol{\xi})$ . The work in [91, 142] relies on a greedy procedure applied to low-fidelity realizations of  $\mathbf{v}$  to identify the set of  $\boldsymbol{\xi}^{(\ell)}$ . The coefficients  $c_{\ell}(\boldsymbol{\xi})$  are computed via an interpolatory condition on the low-fidelity model. The same interpolation rule, i.e., the same coefficients  $c_{\ell}(\boldsymbol{\xi})$ , is then used with the high-fidelity counterpart of the reduced basis snapshots and this results in the bi-fidelity construction (2.1). In this manner, the high-fidelity model is evaluated only at a (small) number of input samples that is equal to the size  $r$  of the reduced basis. Unlike in standard reduced basis techniques this approach does not perform Galerkin (or Petrov-Galerkin) projections of the high-fidelity equations, thus allowing the use of legacy codes in a black-box fashion.

With a different algorithm to construct the reduced basis and the interpolation rule, the work in [27] and [122] examines a similar bi-fidelity approximation on problems involving heat transfer in a ribbed channel and prediction of pressure coefficient over a family of NACA airfoils, respectively. The strategy in [27] was used in [36] to build control variates within a multi-level MC framework.

### 2.1.1 Contribution of This Work

While in practice the above multi-fidelity methods have shown promising results, there is a lack of convenient tools to verify the convergence of the multi-fidelity solution and understand the role of key factors affecting the convergence. These tools must also enable a practitioner to determine, *a priori* and with relatively small cost, whether or not a given low-fidelity model will result in accurate multi-fidelity approximations. To this end, our primary contribution is the

derivation of an error bound that is appropriate for methods as in [91, 142, 27, 36, 122, 71] for bi-fidelity approximation, with a particular emphasis on the role of the rank of the  $\mathbf{v}$  ensemble in the approximation error. This bound is derived specifically for the bi-fidelity procedure in [27, 36, 122, 71] and relies on quantities that are easily estimated with a number of high-fidelity samples that scales favorably with the rank of the  $\mathbf{v}$  ensemble. As a result, the associated error estimate is efficient, easy to compute, and reliably conservative, making it desirable in practical contexts. This method is also presented in its full generality while applied to two examples that demonstrate the potential for practical utility of this error estimate.

The remainder of this manuscript is organized into three main sections. Section 2.2 reviews the mathematical notation and procedure for the bi-fidelity approximation. Section 2.2.4 presents our main result, which focuses on bounding the bi-fidelity approximation error. To show the accuracy of this bound, Section 2.3 provides two numerical experiments and illustrates the utility of the error estimate. Finally, Section 2.4 gives a short summary of the conclusions of this study. In Appendix B, we provide the proof of the main error bound introduced in Section 2.2.4.

## 2.2 Method Detail

As in [91, 142, 27, 36, 122, 71], we consider a bi-fidelity construction of a QoI admitting a low-rank representation of the form (2.1). We presume here that these QoI's are functions defined over the spatial or temporal domain of the problem or can be computed from such functions. While the QoI's could contain the solution over the entire domain of the problem, e.g., temperature over the physical domain, this is often unnecessary and a reduction to some summary of the solution, e.g., temperature along a boundary, can explain the behavior of interest. In a broad sense, this low-rank heuristic implies the existence of a limited number of behaviors for  $\mathbf{v}$  to exhibit, and that for any particular set of inputs  $\boldsymbol{\xi}$ , the corresponding  $\mathbf{v}$  is nearly a linear reconstruction of  $\mathbf{v}$  exhibiting these different traits. In this way, even  $\mathbf{v}$  for possibly highly non-linear problems may be approximated, with some tolerable error, by a linear combination of relatively few basis vectors. This low-rank assumption is actually more than a heuristic: It can be codified via the mathematical notion of  $n$ -

widths [105], and QoI ensembles for differential equations can indeed exhibit this low-rank property [9].

This formulation raises two primary concerns: How do we identify the basis vectors, and how do we identify the appropriate linear combinations of these basis vectors? For computational feasibility, we require solutions to both these concerns that require neither a large number of high-fidelity realizations of  $\mathbf{v}$ , nor modification to the simulation codes.

To present our approach for addressing these concerns, and without loss of generality, we consider the approximation of a collection of  $N$  high-fidelity realizations of the QoI, instead of individual realizations as in (2.1). We denote the  $k$ th realization of the QoI obtained via the high- and low-fidelity models by  $\mathbf{v}_H^{(k)} \in \mathbb{R}^M$  and  $\mathbf{v}_L^{(k)} \in \mathbb{R}^m$ , respectively. These vectors may be of different sizes, but depend on the same parametric input sample  $\boldsymbol{\xi}^{(k)}$ . For the examples presented in Section 2.3, we consider  $\boldsymbol{\xi}$  as a random variable, and  $\mathbf{v}_H^{(k)}$  and  $\mathbf{v}_L^{(k)}$  are drawn in an MC fashion from the joint probability distribution on  $\boldsymbol{\xi}$ . When no probabilistic information on  $\boldsymbol{\xi}$  is available,  $\mathbf{v}_H^{(k)}$  and  $\mathbf{v}_L^{(k)}$  may be generated, e.g., from uniformly distributed samples of  $\boldsymbol{\xi}$ , over some prescribed region of permissible  $\boldsymbol{\xi}$ .

Our analysis relies on arranging  $\{\mathbf{v}_L^{(k)}\}$  and  $\{\mathbf{v}_H^{(k)}\}$  into matrices, denoted  $\mathbf{H} \in \mathbb{R}^{M \times N}$  and  $\mathbf{L} \in \mathbb{R}^{m \times N}$ , to refer to high-fidelity and low-fidelity data, respectively, where the corresponding columns of  $\mathbf{H}$  and  $\mathbf{L}$  are the QoI realizations at the same input samples. That is,  $\mathbf{H}$  and  $\mathbf{L}$  have the same number of columns,  $N$ , but generally have different numbers of rows  $M$  and  $m$ , respectively. The low-rank assumption on the  $\mathbf{v}_H$  ensemble implies that  $\mathbf{H}$  admits a low-rank factorization. We assume the same for  $\mathbf{L}$  but do not require the same rank for  $\mathbf{H}$  and  $\mathbf{L}$ .

**Remark 2.2.1.** *For the interest of presentation and analysis we refer to the full high-fidelity matrix  $\mathbf{H}$ . However, in practice, we require having access to only  $r \ll N$  ( $r$  being the rank of  $\mathbf{v}_H$  as in (2.1)) high-fidelity realizations to generate this bi-fidelity approximation; see Section 2.2.2. To evaluate the bound on the resulting approximation error, we require a number of high-fidelity realizations that is slightly larger than  $r$ , as illustrated in the examples of Section 2.3.*

### 2.2.1 Low-Rank Factorization of Low-Fidelity Data

The method used to identify a reduced basis and an interpolation rule for  $\mathbf{H}$  (or  $\mathbf{v}_H$ ) relies on identifying the same for the low-fidelity data  $\mathbf{L}$ . For this purpose, we form the rank  $r$  *interpolative decomposition* of  $\mathbf{L}$  [23, 84, 57] using the rank-revealing QR factorization [51]. The revealed rank of this factorization is our reduced rank  $r$ , and this provides the rank- $r$  reduced decomposition

$$\mathbf{L}\mathbf{P} \approx \mathbf{Q}_L \mathbf{R}_L, \quad (2.2)$$

for an  $N \times N$  permutation matrix  $\mathbf{P}$ . Then  $\mathbf{Q}_L$  has  $r$  orthonormal columns, and  $\mathbf{R}_L$  has  $r$  rows and is upper triangular. A column partitioning of  $\mathbf{R}_L$  into an  $r \times r$  upper triangular matrix  $\mathbf{R}_L^{(11)}$  and an  $r \times (N - r)$  matrix  $\mathbf{R}_L^{(12)}$ , and replacing  $\mathbf{R}_L^{(12)}$  with the approximation  $\mathbf{R}_L^{(12)} \approx \mathbf{R}_L^{(11)} \mathbf{Z}$  for some  $\mathbf{Z}$  leads to

$$\mathbf{L}\mathbf{P} \approx \mathbf{Q}_L \mathbf{R}_L^{(11)} [\mathbf{I} \mid \mathbf{Z}], \quad (2.3)$$

where  $\mathbf{I}$  is the  $r \times r$  identity matrix. When  $\mathbf{R}_L^{(11)}$  is ill-conditioned, [23] suggests a solution  $\mathbf{Z}$  with minimum  $\|\mathbf{Z}\|_F$ , where  $\|\cdot\|_F$  denotes the Frobenius norm. The rank  $r$  factorization (2.3) can be rewritten as

$$\begin{aligned} \mathbf{L} &\approx \mathbf{L}(r) [\mathbf{I} \mid \mathbf{Z}] \mathbf{P}^T; \\ &= \mathbf{L}(r) \mathbf{C}_L; \end{aligned} \quad (2.4)$$

$$:= \hat{\mathbf{L}}, \quad (2.5)$$

where  $\mathbf{L}(r) = \mathbf{Q}_L \mathbf{R}_L^{(11)}$  – referred to as the *column skeleton* of  $\mathbf{L}$  [23] – contains the first  $r$  columns of  $\mathbf{L}\mathbf{P}$  and  $\mathbf{C}_L := [\mathbf{I} \mid \mathbf{Z}] \mathbf{P}^T$ . In words, (2.4) identifies a set of  $r$  columns of  $\mathbf{L}$ , i.e., a reduced basis for  $\mathbf{L}$ , along with a coefficient matrix  $\mathbf{C}_L$  specifying the linear combination of the basis vectors to approximate all columns in  $\mathbf{L}$ . Hence the name *interpolative decomposition*. If we identify the entries of  $\mathbf{C}_L$  with the coefficients  $c_\ell(\boldsymbol{\xi})$  in (2.1), then (2.4) is the matrix formulation of (2.1) for  $N$  low-fidelity realizations  $\{\mathbf{v}_L^{(k)}\}$ .

For the sake of completeness, we next report some properties of the factorization (2.4) from [84], which we will use later. For detailed analyses of the interpolative decomposition, the interested reader is referred to [23, 84, 57].

**Lemma 2.2.1.** *(Lemma 3.1 of [84].) Let  $\|\cdot\|$  denote the matrix induced  $\ell_2$  norm. For any positive integer  $r \leq \min\{m, N\}$ ,*

$$(1) \quad \|\mathbf{C}_L\| \leq \sqrt{r(N-r)+1},$$

$$(2) \quad \mathbf{L} = \hat{\mathbf{L}}, \text{ when } r = m \text{ or } r = N, \text{ and}$$

$$(3) \quad \|\mathbf{L} - \hat{\mathbf{L}}\| \leq \sqrt{r(N-r)+1} \sigma_{r+1} \text{ when } r < \min\{m, N\}, \text{ where } \sigma_{r+1} \text{ is the } (r+1)\text{st greatest singular value of } \mathbf{L}.$$

**Remark 2.2.2.** *In practice, the approximation rank  $r$  is not known a priori and thus the pivoted Gram-Schmidt step involved in the QR factorization (2.2) is continued with larger ranks until  $\|\mathbf{L} - \hat{\mathbf{L}}\|$  reaches a specified accuracy.*

**Remark 2.2.3.** *Although not investigated here, other factorizations such as those in [80, 33, 30, 103] are equally applicable to produce an interpolative decomposition of  $\mathbf{L}$ .*

## 2.2.2 High-Fidelity Approximation via Basis Update and Low-Fidelity Interpolation

In addition to a rank  $r$  factorization of  $\mathbf{L}$ , the permutation matrix  $\mathbf{P}$  in the interpolative decomposition (2.4) identifies the set of  $r$  input samples  $\boldsymbol{\xi}^{(k)}$  corresponding to the basis vectors in  $\mathbf{L}(r)$ . This identification is one primary reason for employing an interpolative decomposition instead of the more widely-used and  $\ell_2$ -optimal singular value decomposition (SVD) of  $\mathbf{L}$ . As in [91, 142, 27, 36, 122, 71], our approach to approximate  $\mathbf{H}$  is to replace  $\mathbf{L}(r)$  with  $\mathbf{H}(r)$ , a matrix of vectors from the high-fidelity model corresponding to the same input samples  $\boldsymbol{\xi}^{(k)}$  and with the same arrangement as in  $\mathbf{L}(r)$ . Stated differently, we consider the high-fidelity counterpart of  $\mathbf{L}(r)$

as the reduced basis for  $\mathbf{H}$  (or  $\mathbf{v}_H^{(k)}$ ). In this case we set the bi-fidelity, rank  $r$  approximation of  $\mathbf{H}$ , denoted by  $\hat{\mathbf{H}}$ , to

$$\hat{\mathbf{H}} := \mathbf{H}(r)\mathbf{C}_L, \quad (2.6)$$

where  $\mathbf{C}_L$  is the coefficient matrix computed from the low-fidelity approximation (2.4). In other words, the interpolation rule learned from the low-fidelity model is applied to approximate the high-fidelity realizations. In the case that we have a construction of this form, where the difference between  $\hat{\mathbf{L}}$  and  $\hat{\mathbf{H}}$  is limited to the changes only in the basis vectors associated with the low- and high-fidelity models, we say that  $\hat{\mathbf{L}}$  is *lifted* to the approximation  $\hat{\mathbf{H}}$ . For this paper we say that  $\hat{\mathbf{L}}$  and  $\hat{\mathbf{H}}$  are *corresponding estimates* (of  $\mathbf{L}$  and  $\mathbf{H}$ , respectively). Our analysis presented next applies to any such corresponding estimates.

**Remark 2.2.4** (Utility of  $\hat{\mathbf{H}}$  for UQ). *While outside the scope of the present study, the bi-fidelity realizations  $\hat{\mathbf{v}}_H^{(k)}$ , i.e., the columns of  $\hat{\mathbf{H}}$ , may be used in place of  $\mathbf{v}_H^{(k)}$  in methods such as MC simulation or its variants, sparse grid stochastic collocation [135], non-intrusive PCE [67, 29, 25, 55], etc., to approximate the statistics of  $\mathbf{v}_H$  or perform sensitivity analysis. The accuracy of such estimates, however, depends on the accuracy of the bi-fidelity estimates  $\hat{\mathbf{v}}_H^{(k)}$ .*

### 2.2.3 Number of Low- and High-Fidelity Samples

Notice that the bi-fidelity construction (2.6) relies on  $r$  high-fidelity samples. Therefore, when  $\mathbf{v}_H$  admits a low-rank representation, i.e.,  $r$  is small, only a small number of high-fidelity samples are needed. To generate a rank- $r$  approximation of  $\mathbf{L}$ , a theoretical lower bound on the number of required low-fidelity model evaluations  $N$  (and  $m$ ) is  $r$ . In practice, however, a considerably larger number of low-fidelity samples may be needed to identify a representative reduced basis. Finding an optimal  $N$  to reduce the cost of low-fidelity model evaluations depends on the problem at hand, and on how  $\mathbf{v}_L$  varies with respect to  $\boldsymbol{\xi}$ . For scenarios where the low-fidelity model is not substantially cheaper than its high-fidelity counterpart, adaptive sampling strategies – similar to

those in [20, 68] – can be employed to reduce the number of low-fidelity model evaluations. This is, however, outside the scope of this manuscript.

#### 2.2.4 Convergence Analysis and Its Practical Application

We are now prepared to state the main result that bounds the bi-fidelity error,  $\|\hat{\mathbf{H}} - \mathbf{H}\|$ , where  $\|\cdot\|$  is matrix induced  $\ell_2$  norm. However, we first present a high-level discussion on our analysis approach. The core of our convergence results rely on the assumption that there exists a matrix  $\mathbf{T}$ , with  $\|\mathbf{T}\|$  bounded, so that  $\mathbf{H} \approx \mathbf{T}\mathbf{L}$ . The difference in these matrices is denoted  $\mathbf{E}$ , and we require that  $\|\mathbf{E}\|$  is small. Restated, our bound involves the matrix relation given by

$$\mathbf{H} = \mathbf{T}\mathbf{L} + \mathbf{E}. \quad (2.7)$$

The matrix  $\mathbf{T}$  is essentially a lifting operator from low-fidelity discretization space to high-fidelity discretization space. Our theory outlined in Appendix B computes this matrix explicitly, but does not require any knowledge about the discretizations or assumptions between the low- and high-fidelity models. We emphasize that  $\mathbf{T}$  is only needed for theoretical analysis and our error bounds do not require its construction.

It is straightforward to verify that  $\|\hat{\mathbf{H}} - \mathbf{H}\|$  can be bounded as

$$\begin{aligned} \|\mathbf{H} - \hat{\mathbf{H}}\| &\leq \|\mathbf{H} - \mathbf{T}\mathbf{L}\| + \|\mathbf{T}\mathbf{L} - \mathbf{T}\hat{\mathbf{L}}\| + \|\mathbf{T}\hat{\mathbf{L}} - \hat{\mathbf{H}}\|; \\ &\leq \|\mathbf{H} - \mathbf{T}\mathbf{L}\| + \|\mathbf{T}\|\|\mathbf{L} - \hat{\mathbf{L}}\| + \|\mathbf{T}\mathbf{L} - \mathbf{H}\|\|\mathbf{C}_L\|; \\ &\leq (1 + \|\mathbf{C}_L\|)\|\mathbf{E}\| + \|\mathbf{T}\|\|\mathbf{L} - \hat{\mathbf{L}}\|. \end{aligned} \quad (2.8)$$

When  $\|\mathbf{T}\|$  is bounded and  $\|\mathbf{E}\|$  is small, (2.8) suggests an accurate bi-fidelity estimate as  $\|\mathbf{C}_L\|$  is bounded, and  $\|\mathbf{L} - \hat{\mathbf{L}}\|$  is small given that  $\mathbf{L}$  is low-rank; see Lemma 2.2.1. In general, there is no guarantee of the existence of a mapping  $\mathbf{T}$  of  $\mathbf{L}$  to  $\mathbf{H}$  that gives an  $\mathbf{E}$  with small norm. We therefore seek to establish a condition on  $\mathbf{L}$  (in relation to  $\mathbf{H}$ ) to ensure the existence of such  $\mathbf{T}$  and  $\mathbf{E}$  matrices.

For a finite  $\tau \geq 0$ , define

$$\epsilon(\tau) := \lambda_{\max}(\mathbf{H}^T \mathbf{H} - \tau \mathbf{L}^T \mathbf{L}), \quad (2.9)$$

where  $\lambda_{\max}(\cdot)$  denotes the largest eigenvalue of a matrix, and  $\mathbf{H}^T \mathbf{H}$  and  $\mathbf{L}^T \mathbf{L}$  are the Gramians of the high- and low-fidelity matrices, respectively. We describe in Appendix B that if  $\epsilon(\tau)$  is small enough then matrices  $\mathbf{T}$  and  $\mathbf{E}$  with our desired properties can be constructed.

For a fixed  $\tau$ ,  $\epsilon(\tau)$  has various interpretations. It is the smallest  $\epsilon$  such that

$$\epsilon \mathbf{I} + \tau \mathbf{L}^T \mathbf{L} - \mathbf{H}^T \mathbf{H} \quad (2.10)$$

is a positive semi-definite matrix. Second,  $\epsilon(\tau)$  has a geometrical interpretation that demonstrates its significance in this context of bi-fidelity approximation. The above equation implies that  $\epsilon(\tau)$  is the smallest value such that for any  $\mathbf{x} \in \mathbb{R}^N$ ,

$$\|\mathbf{H}\mathbf{x}\|^2 \leq \tau \|\mathbf{L}\mathbf{x}\|^2 + \epsilon(\tau) \|\mathbf{x}\|^2.$$

This guarantees that for every  $\mathbf{x}$ ,  $\sqrt{\tau} \mathbf{L}\mathbf{x}$  is further from the origin than  $\mathbf{H}\mathbf{x}$ , to within a margin governed by  $\epsilon(\tau)$  and  $\|\mathbf{x}\|$ . This ensures that every realized  $\mathbf{H}\mathbf{x}$  can be reached by rotating  $\mathbf{L}\mathbf{x}$  through the application of a scaling related to  $\tau$ , and adding a small correction which is bounded based on  $\epsilon(\tau)$  and  $\|\mathbf{x}\|$ . The relationship between these interpretations is further discussed in Appendix B, and we note this rotation, scaling, and additive correction is essentially utilized to prove the following theorem completing the bound in (2.8).

**Theorem 2.2.1.** *For any  $\tau \geq 0$ , let  $\epsilon(\tau)$  be as in (2.9). Let  $\hat{\mathbf{H}}$  and  $\hat{\mathbf{L}}$  be corresponding estimates of rank  $r$  with coefficients  $\mathbf{C}_L$ , and let  $\sigma_k$  denote the  $k$ th largest singular value of  $\mathbf{L}$ . Then,*

$$\|\mathbf{H} - \hat{\mathbf{H}}\| \leq \min_{\tau, k \leq \text{rank}(\mathbf{L})} \rho_k(\tau) \quad (2.11a)$$

$$\rho_k(\tau) := \left[ (1 + \|\mathbf{C}_L\|) \sqrt{\tau \sigma_{k+1}^2 + \epsilon(\tau)} + \|\mathbf{L} - \hat{\mathbf{L}}\| \sqrt{\tau + \epsilon(\tau) \sigma_k^{-2}} \right]. \quad (2.11b)$$

When  $k = \text{rank}(\mathbf{L})$ , we set  $\sigma_{k+1} = 0$ .

The proof of this theorem is given in Appendix B; here we provide insight and implications of its result.



An actual evaluation of the bound requires estimating  $\tau$  and  $\epsilon(\tau)$ , which can be achieved by minimizing the bound over a range of  $k$  and  $\tau$ ; see Algorithm 1. A naïve attempt to estimate  $\epsilon(\tau)$  from (2.9) seems to require the entire matrix  $\mathbf{H}$ . However, we propose in Section 2.2.5 a procedure that uses only a relatively small number of high- and low-fidelity realizations. Notice that aside from  $\epsilon(\tau)$ , the bound (2.11) has no dependence on the high-fidelity model. Let

$$B_1 = (1 + \|\mathbf{C}_L\|) \sqrt{\tau \sigma_{k+1}^2 + \epsilon(\tau)}; \quad (2.12)$$

$$B_2 = \|\mathbf{L} - \hat{\mathbf{L}}\| \sqrt{\tau + \epsilon(\tau) \sigma_k^{-2}}. \quad (2.13)$$

If the approximation is accurate, the expressions in (2.12) and (2.13) will be small. Specifically, observe that  $B_1$  is small when both  $\tau \sigma_{k+1}^2 \ll 1$  and the optimal  $\epsilon(\tau)$  is small. Recall  $\|\mathbf{C}_L\| \leq \sqrt{r(N-r)+1}$  from Lemma 2.2.1. When, for a given  $k$ ,  $\sigma_{k+1}$  is small, i.e., the singular values of  $\mathbf{L}$  decay quickly, the magnitude of  $B_1$  is governed by the optimal  $\epsilon(\tau)$ . To see that  $B_2$  is small, note that if the low-rank reconstruction is accurate for the low-fidelity model, then  $\|\mathbf{L} - \hat{\mathbf{L}}\| \leq \sqrt{r(N-r)+1} \sigma_{r+1}$  is small following Lemma 2.2.1. Given this,  $B_2$  is also small when  $\sigma_{k+1}/\sigma_k \ll 1$ ,  $\tau \sigma_{k+1}^2 \ll 1$ , and the optimal  $\epsilon(\tau)$  is small. We note that the requirements  $\sigma_{k+1}/\sigma_k \ll 1$  and  $\tau \sigma_{k+1}^2 \ll 1$  highlight the significance of the low-rank assumption on  $\mathbf{L}$  in the success of this bi-fidelity construction. Specifically, a low-rank approximation will have a  $k$  such that  $\sigma_{k+1}$  is significantly smaller than  $\sigma_k$ .

Using Lemma 2.2.1, the bound (2.11) can be simplified to

$$\|\mathbf{H} - \hat{\mathbf{H}}\| \leq \min_{\tau, k \leq \text{rank}(\mathbf{L})} \left[ \sqrt{r(N-r)+1} \left( 1 + \frac{\sigma_{k+1}}{\sigma_k} \right) \sqrt{\tau \sigma_k^2 + \epsilon(\tau)} \right].$$

In our numerical experiments, we use the sharper error estimate in (2.11).

We now address the computation of an optimal  $\epsilon(\tau)$ , or, more precisely an optimal pair  $(\tau, \epsilon(\tau))$ . Specifically we consider an estimate  $\hat{\epsilon}(\tau)$  with relatively few evaluations of the high-fidelity model. Such an estimate allows us to approximate the bound (2.11) on  $\|\mathbf{H} - \hat{\mathbf{H}}\|$ , which measures the quality of the bi-fidelity approximation.

### 2.2.5 Estimating $\epsilon(\tau)$ Using Limited High-Fidelity Data

The definition of  $\epsilon(\tau)$  in (2.9) depends on full realizations of the high-fidelity model which is untenable in many practical situations and its direct computation would significantly detract from the utility of the bi-fidelity error estimate. Consider Gramian matrices and an estimate  $\hat{\epsilon}$  defined, for a normalizing constant  $c$ , by

$$\mathbf{G}_H := \mathbf{H}^T \mathbf{H}; \quad (2.14)$$

$$\mathbf{G}_L := \mathbf{L}^T \mathbf{L}; \quad (2.15)$$

$$\hat{\epsilon}(\tau) := c \lambda_{\max}(\hat{\mathbf{G}}_H - \tau \hat{\mathbf{G}}_L), \quad (2.16)$$

for estimates  $\hat{\mathbf{G}}_H$  and  $\hat{\mathbf{G}}_L$  of  $\mathbf{G}_H$  and  $\mathbf{G}_L$ , respectively. We note the similarity of (2.9) and (2.16), where the latter makes use of quantities with the hat notation. Of key practical note is that  $\mathbf{G}_H$  inherits the low rank nature of  $\mathbf{H}$ , so that the estimate of  $\epsilon(\tau)$  can be accurate with a small number of samples. Recall that the number of columns of  $\mathbf{H}$  and  $\mathbf{L}$  is  $N$ . Let  $n$  denote a small number of columns that are sub-sampled from  $\mathbf{H}$  and  $\mathbf{L}$ , and  $\hat{\mathbf{G}}_H$  and  $\hat{\mathbf{G}}_L$  are the Gramian matrices associated with these  $n$  columns. We set  $c = N/n$  as the normalizing constant throughout the remainder of this work. With these Gramian estimates, we can construct  $\hat{\epsilon}(\tau)$  for any  $\tau$  using (2.16). To approximate the bound in (2.11), we may replace  $\epsilon(\tau)$  with  $\hat{\epsilon}(\tau)$ , and then identify the minimum achieved value over a range of  $k$  and  $\tau$  values, noting that the search over  $k$  requires only additionally knowing the appropriate singular values of  $\mathbf{L}$ . This requires  $n$  high-fidelity samples (to form the Gramian matrix  $\hat{\mathbf{G}}_H$ ). Our numerical results in Section 2.3 show empirically that  $n$  slightly larger than  $r$  is sufficient to estimate the optimal pair  $(\tau, \epsilon(\tau))$ . Notice, however, that the low-fidelity error  $\|\mathbf{L} - \hat{\mathbf{L}}\|$  in (2.11) is calculated using all  $N$  low-fidelity samples.

Algorithm 1 summarizes our proposed approach to evaluate the bound (2.11) as an estimate for the bi-fidelity error  $\|\mathbf{H} - \hat{\mathbf{H}}\|$ .

**Algorithm 1:** Algorithm for estimating the error bound (2.11).

- (1) Initialize a vector of values for  $\tau \geq 0$ .
- (2) Use  $n$  high- and low-fidelity realizations to set the Gramian matrices  $\hat{\mathbf{G}}_L$  and  $\hat{\mathbf{G}}_H$  as explained in Section 2.2.5.
- (3) For all values of  $\tau$  evaluate the corresponding  $\hat{\epsilon}(\tau)$  using (2.16).
- (4) For each  $k \in \{0, \dots, \text{rank}(\mathbf{L})\}$  evaluate (2.11) for all pairs  $(\tau, \hat{\epsilon}(\tau))$  computed in Step 3 above.
- (5) Choose the minimum value achieved by (2.11) over all  $k$  and  $(\tau, \hat{\epsilon}(\tau))$  as an estimate for  $\|\mathbf{H} - \hat{\mathbf{H}}\|$ .

## 2.3 Numerical Examples

To investigate various aspects of the proposed bi-fidelity approximation and the associated error estimate, we consider two practically motivated problems.

### 2.3.1 Test Case 1: Heat Driven Cavity Flow

A practical case for consideration comes from temperature-driven fluid flow in a cavity [81, 82, 108, 100, 60, 59, 36, 61], where the QoI is the steady state heat flux along the hot wall; see Figure 2.1. The left vertical wall has a random temperature  $T_h$ , referred to as the hot wall, while the right vertical wall, referred to as the cold wall, has a spatially varying stochastic temperature  $T_c < T_h$  with constant mean  $\bar{T}_c$ . Both horizontal walls are adiabatic. The reference temperature and the reference temperature difference are defined as  $T_{ref} = \bar{T}_c$  and  $\Delta T_{ref} = T_h - \bar{T}_c$ , respectively. Under small temperature difference assumption, i.e., the Boussinesq approximation, the normalized governing equations are given by, [81],

$$\begin{aligned}
 \frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} &= -\nabla p + \frac{\text{Pr}}{\sqrt{\text{Ra}}} \nabla^2 \mathbf{u} + \text{Pr} \Theta \mathbf{e}_y, \\
 \nabla \cdot \mathbf{u} &= 0, \\
 \frac{\partial \Theta}{\partial t} + \nabla \cdot (\mathbf{u} \Theta) &= \frac{1}{\sqrt{\text{Ra}}} \nabla^2 \Theta,
 \end{aligned} \tag{2.17}$$

where  $\mathbf{e}_y$  is the unit vector  $(0, 1)$ ,  $\mathbf{u} = (u, v)$  is velocity vector field,  $\Theta = (T - T_{ref})/\Delta T_{ref}$  is normalized temperature,  $p$  is pressure, and  $t$  is time. Zero velocity boundary conditions on all walls (in both directions) are assumed. For more details on the normalization of the variables in (2.17), we refer the interested reader to [108, 81]. Prandtl and Rayleigh numbers are defined, respectively, as  $\text{Pr} = \nu/\alpha$  and  $\text{Ra} = g\tau\Delta T_{ref}L^3/(\nu\alpha)$ . Specifically,  $L$  is the length of the cavity,  $g$  is gravitational acceleration,  $\nu$  is kinematic viscosity,  $\alpha$  is thermal diffusivity, and the coefficient of thermal expansion is given by  $\tau$ . In this example, we set  $g = 10$ ,  $L = 1$ ,  $\tau = 0.5$ , and  $\text{Pr} = 0.71$ . We use a finite volume method for the discretization of (2.17).

### 2.3.1.1 Sources of Uncertainty

On the cold wall, a temperature distribution with stochastic fluctuations is applied as

$$T(x = 1, y) = \bar{T}_c + T'(y), \quad (2.18)$$

where  $\bar{T}_c = 100$  is a constant expected temperature. The fluctuation  $T'(y)$  is given by the truncated Karhunen-Loève-like expansion

$$T'(y) = \sigma_T \sum_{i=1}^{d_T} \sqrt{\lambda_i} \varphi_i(y) \xi_i, \quad (2.19)$$

where  $d_T = 50$  and  $\sigma_T = 2$ . Here, each  $\xi_i$  is assumed to be an independent and identically distributed uniform random variable on  $[-1, 1]$ , and  $\{\lambda_i\}_{i=1}^{d_T}$  and  $\{\phi_i(y)\}_{i=1}^{d_T}$  are the  $d_T$  largest eigenvalues and the corresponding eigenfunctions of the exponential covariance kernel

$$C_{TT}(y_1, y_2) = \exp\left(-\frac{|y_1 - y_2|}{l_T}\right), \quad (2.20)$$

with correlation length  $l_T = 0.15$ . An example of the cold boundary condition is shown in Figure 2.1. The temperature on the hot wall  $T_h$  is also assumed to be random and uniformly distributed over [105, 109]. Finally, we consider the viscosity  $\nu$  to be uniformly distributed over [0.004, 0.01]. In total, the dimension of the random input  $\boldsymbol{\xi} = (\xi_1, \dots, \xi_{50}, T_h, \nu)$  is 52. The QoI, heat flux along the hot wall, is represented by a vector of heat flux values at the grid points (along the hot wall) corresponding the high- or low-fidelity meshes.

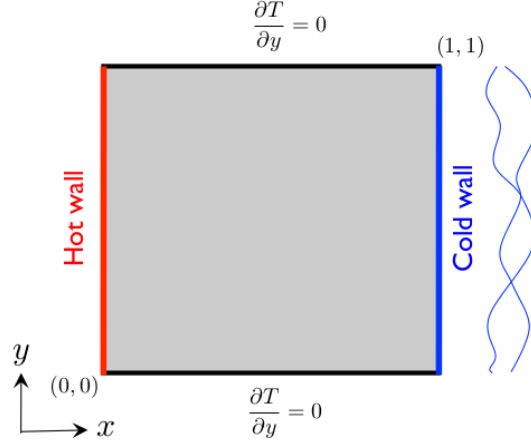


Figure 2.1: Schematic of the heat driven cavity flow problem, reproduced from Figure 5 of [36].

### 2.3.1.2 Bi-Fidelity Approximation and Error Bound Estimate

To compute realizations of the QoI, we use finite volume discretizations of various resolutions to solve (2.17), where the highest resolution, a  $256 \times 256$  grid point mesh, is used as the high-fidelity model, while relatively coarse meshes of  $128 \times 128$ ,  $64 \times 64$ ,  $32 \times 32$ , and  $16 \times 16$  grid points are used for the corresponding low-fidelity model. All meshes are spatially uniform. We consider bi-fidelity approximations of various ranks  $r$ , corresponding to the number of basis vectors in the column skeleton matrices  $\mathbf{H}(r)$  and  $\mathbf{L}(r)$  introduced in Section 2.2.4. We also consider bounds derived from computations of optimal pair  $(\tau, \hat{\epsilon}(\tau))$  as explained in Section 2.2.5. These estimates are drawn from  $n$  randomly generated high-fidelity samples, i.e.,  $n$  randomly selected columns of  $\mathbf{H}$ , and we consider estimates computed from various sample sizes  $n$ .

Figure 2.2 displays realizations of heat flux along the hot wall for a random sample of  $\xi$ , obtained by the high-fidelity, low-fidelity, and rank  $r = 10$  bi-fidelity models. We observe the close agreement between the bi- and high-fidelity solutions even when the  $16 \times 16$  mesh is used as the low-fidelity model. Figure 2.3 provides four histograms of the low- and bi-fidelity error to gauge the performance of all 100 QoI realizations. For (a)-(d), which use increasingly refined meshes for the low-fidelity model, we see an improvement of the bi-fidelity performance over that of the associated low-fidelity model.

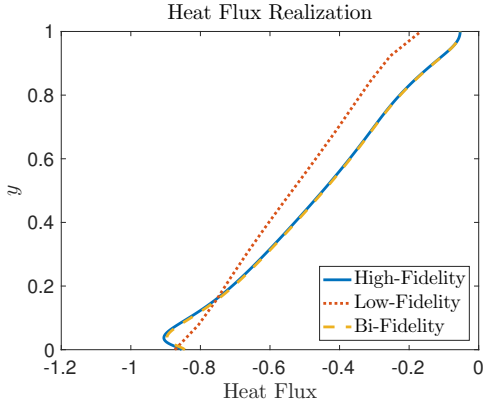
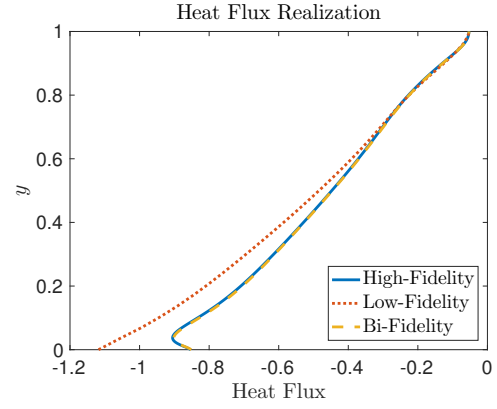
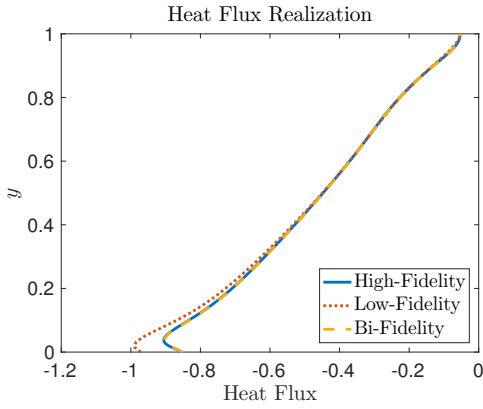
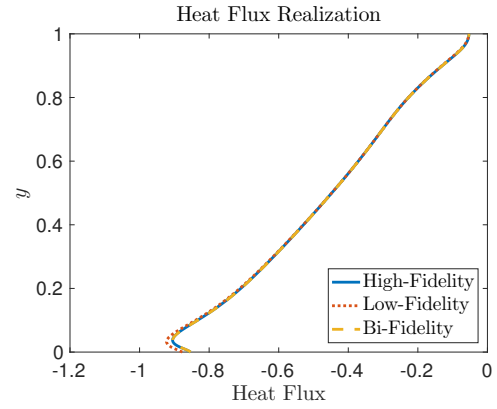
(a)  $16 \times 16$  Low-Fidelity Mesh(b)  $32 \times 32$  Low-Fidelity Mesh(c)  $64 \times 64$  Low-Fidelity Mesh(d)  $128 \times 128$  Low-Fidelity Mesh

Figure 2.2: Realizations of heat flux along the hot wall for a randomly selected input  $\xi$ . Shown are the low-fidelity, high-fidelity, and rank  $r = 10$  bi-fidelity estimates for the various low-fidelity models.

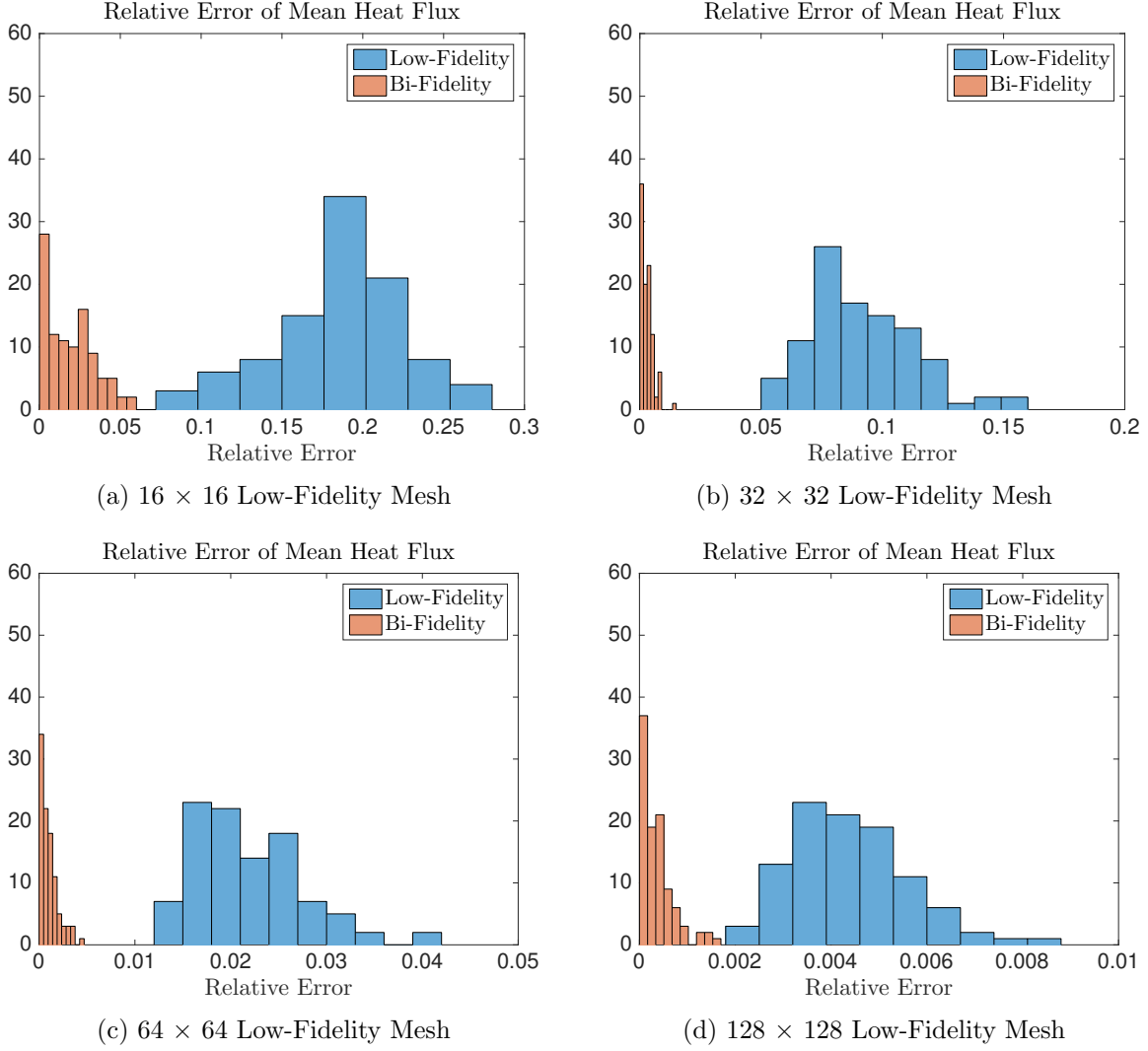


Figure 2.3: Histograms of low-fidelity and rank  $r = 10$  bi-fidelity errors normalized by  $\|\mathbf{H}\|$  for various low-fidelity models from 100 data samples.

We next consider the error bound efficacy of using  $\hat{\epsilon}(\tau)$  from (2.16), in place of  $\epsilon(\tau)$  from (2.9), in Theorem 2.2.1. This efficacy is defined as the ratio of the error estimate over the actual error in the approximation of  $\mathbf{H}$  so that an efficacy of 1 implies the error estimate is exact, and an efficacy greater than 1 implies that the computed error estimate is in fact a bound. Figure 2.4 shows this error efficacy for various low-fidelity models, approximation ranks  $r$ , and sample sizes  $n$ . For each pair  $(n, r)$  the reported error ratio is the average of 30 error ratios, each computed from independent sets of  $n$  high-fidelity samples. We note that the error efficacy is greater than one whenever the sample size  $n$  is greater than the approximation rank  $r$ , and that, as long as this condition holds, the error efficacy is less than ten, implying that the error bound is off by less than one order of magnitude. This is of great utility in determining either a sample size  $n$  for a given approximation rank  $r$  or, conversely, determining an admissible rank of approximation for a given sample size for which the computed error estimate is still valid. These estimates indicate that the error bounds are generally conservative with relatively few samples.

To address the estimation of (2.11) numerically, we consider in Figure 2.5 how  $\hat{\epsilon}(\tau)$  in (2.16) behaves as a function of  $\tau$  for four combinations of rank  $r$  and sample size  $n$ . We mark points that optimize (2.11) with the corresponding bi-fidelity rank  $r$ , and note the reduction in the optimal values of  $\tau$  and  $\hat{\epsilon}(\tau)$  as the mesh of the low-fidelity model is refined. Figure 2.5 (a) and (b), which provide values of  $\hat{\epsilon}(\tau)$  for  $r = 5$ , show that while estimates of  $\hat{\epsilon}(\tau)$  differ for the two values of  $n$ , points of  $(\tau, \hat{\epsilon}(\tau))$  that optimize (2.11) are close. In contrast, Figure 2.5 (c) and (d), which provide the the same set of results but for  $r = 10$ , show dissimilar optimal points. These differences correspond to the ability to describe points of  $(\tau, \epsilon(\tau))$ . Recall, from the results of Figure 2.4 that small values of  $n$  can result in underestimating the bound. For the scenario with  $r = 10$  and  $n = 10$ , these error bound estimates are not reliable, which is further highlighted by optimal values occurring at  $\hat{\epsilon}(\tau) = 0$  in Figure 2.5 (c).

We can also explore the behavior of  $\rho_k(\tau)$  as a function of  $\tau$  in (2.11b), noting that minimizing over  $\tau$  and  $k$  yields our main estimate (2.11a). This function is seen for different low-fidelity model meshes in Figure 2.6. We observe that as the low-fidelity meshes are refined,  $\tau$  converges to 1,



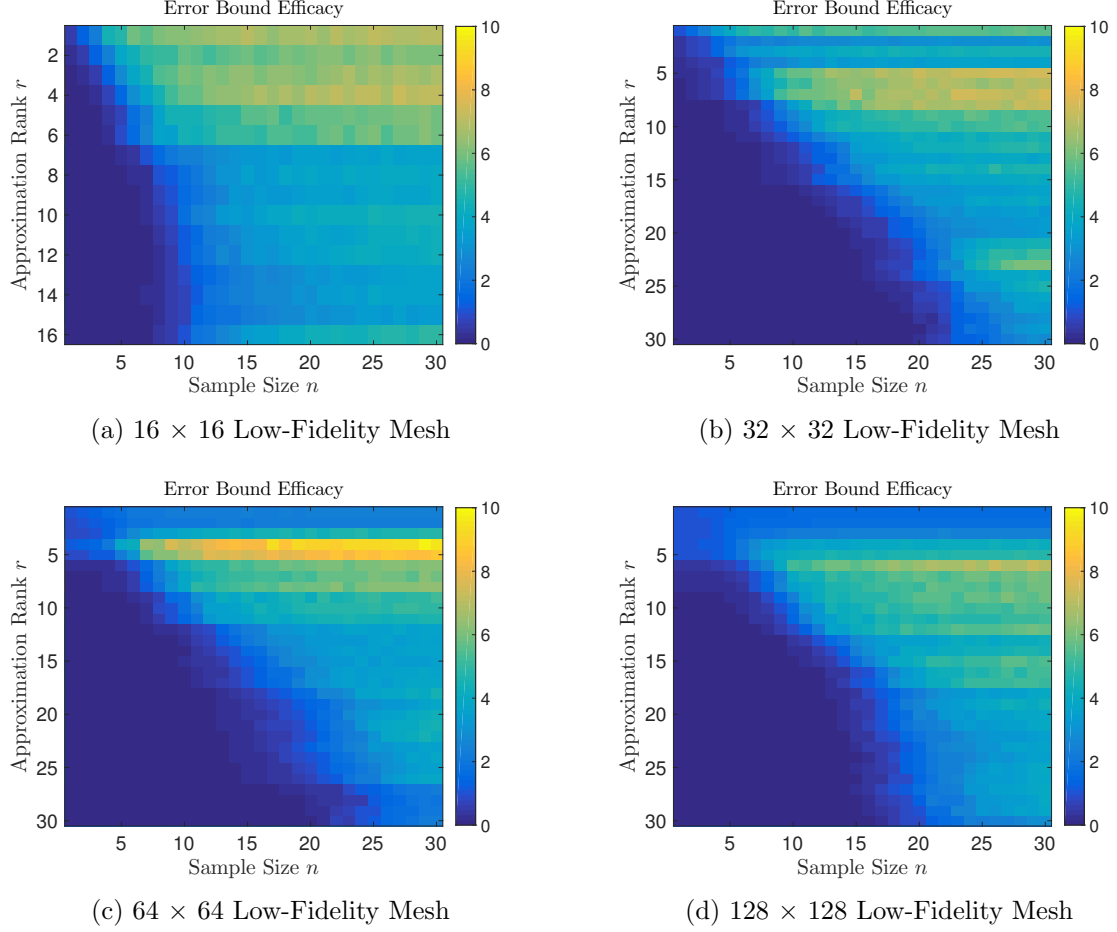


Figure 2.4: Identification of error bound efficacy, i.e., the (average) ratio between the error estimated from (2.11) and the true error, for various low-fidelity meshes, approximation ranks  $r$ , and sample sizes  $n$ . For each pair  $(n, r)$  the reported error ratio is the average of 30 ratios, each computed from an independent set of  $n$  high- and low-fidelity samples via Algorithm 1.

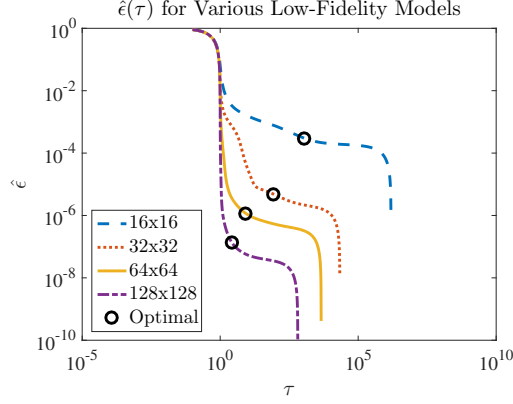
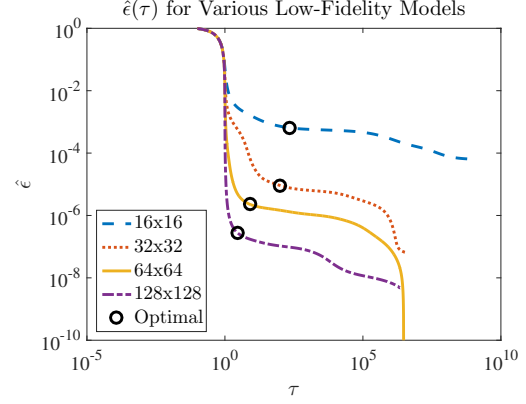
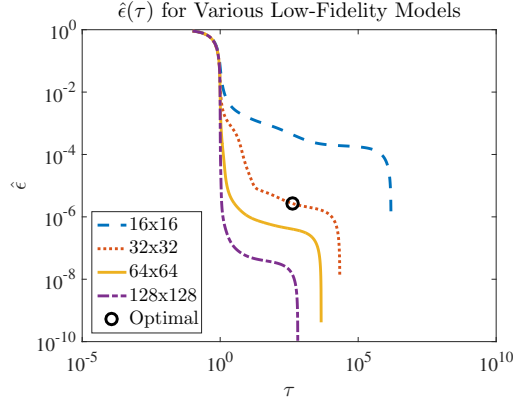
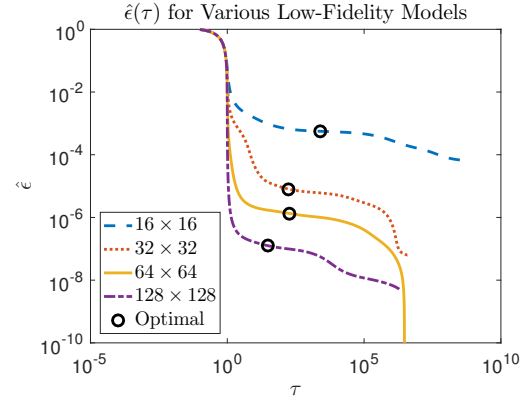
(a) Rank  $r = 5$ , Sample Size  $n = 10$ (b) Rank  $r = 5$ , Sample Size  $n = 20$ (c) Rank  $r = 10$ , Sample Size  $n = 10$ (d) Rank  $r = 10$ , Sample Size  $n = 20$ 

Figure 2.5:  $\hat{\epsilon}(\tau)$  from (2.16) as a function of  $\tau$  for different low-fidelity model meshes with four combinations of bi-fidelity ranks  $r = 5, 10$  and sample sizes  $n = 10, 20$ . Values which optimize (2.11) for a rank  $r$  bi-fidelity approximation are marked by black circles. Curves without optimal value markers displayed identify optimal bounds at  $\hat{\epsilon}(\tau) = 0$ .

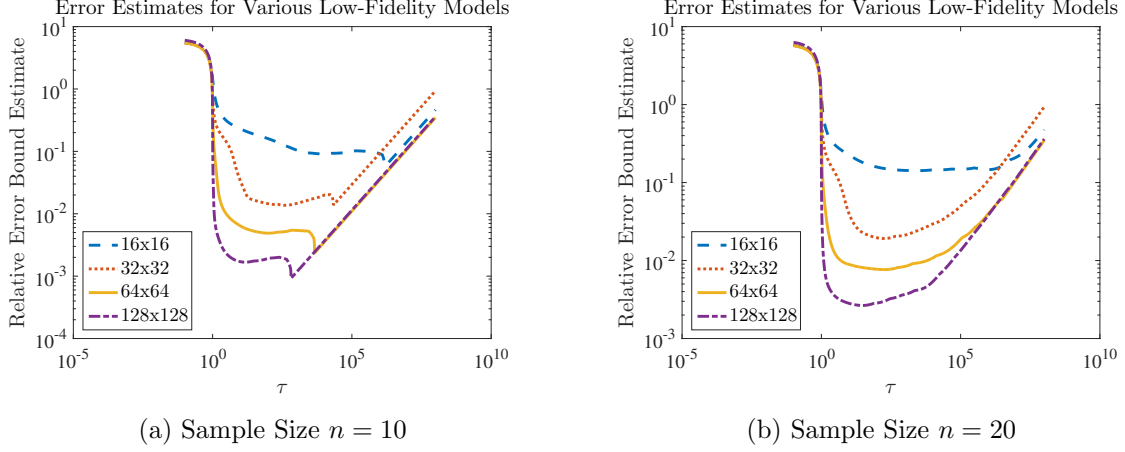


Figure 2.6: Minimum (over  $k$ ) of error function  $\rho_k(\tau)$  in (2.11b) for various low-fidelity models using an approximation rank  $r = 10$ . In (a), minimizations occur (from coarsest to finest mesh) at  $k = 13, 8, 9$ , and  $85$ . In (b), minimizations occur (from coarsest to finest mesh) at  $k = 7, 8, 9$ , and  $9$ . Errors are normalized by  $\|\mathbf{H}\|$ .

corresponding to a broadly defined convergence of low-fidelity models to the high-fidelity model. We also note, that there is not a high level of sensitivity with regards to  $\tau$ , and that a large range of  $\tau$  produces comparable bounds. Figure 2.6 also illustrates the benefit that the optimization of  $\tau$  need not be particularly accurate, and that the general shape of these curves is smooth enough that derivative-based optimization can be employed, limiting the number of  $\tau$  for which  $\hat{e}(\tau)$  needs to be evaluated.

It is worth highlighting that the heuristic employed here is that the size of the matrices used in  $\hat{e}(\tau)$  need only depend on the approximation rank  $r$  and scale nearly as such. As a result, the cost of evaluating  $\hat{e}(\tau)$  will typically not be prohibitively large, as illustrated in Figure 2.7 by the rapid convergence of the optimal  $\hat{e}(\tau)$  with respect to  $n$ . Similar to Figure 2.4, this figure also shows that when the sample size is too small (relative to  $r$ ), e.g.,  $n = 8$ , the error estimate is below the actual error. Finally, Figure 2.8 shows the convergence of the  $r = 10$  bi-fidelity solution as a function of the low-fidelity mesh size. The error bound estimate – relative to  $\mathbf{H}$  – is displayed for two sample sizes of  $n$ , and the low-fidelity error – also relative to  $\mathbf{H}$  – is provided for comparison. Notice that, aside for a single point from the coarsest low-fidelity model, all the estimated errors

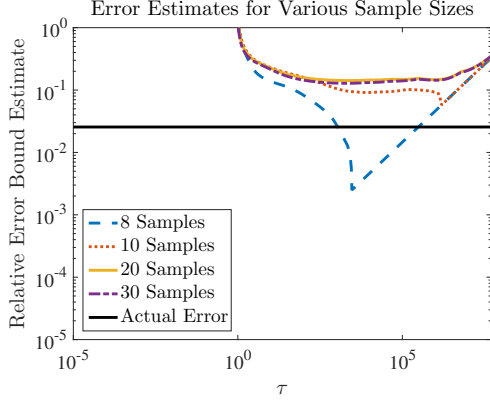
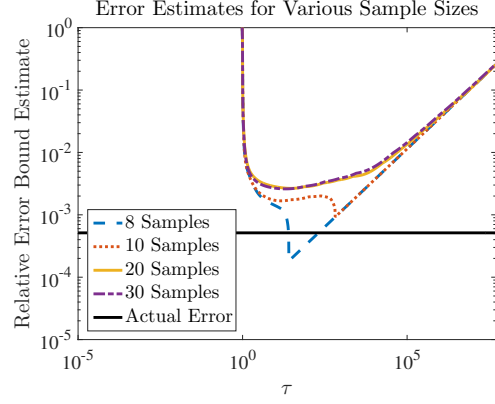
(a)  $16 \times 16$  Low-Fidelity Mesh(b)  $128 \times 128$  Low-Fidelity Mesh

Figure 2.7: Minimum (over  $k$ ) of error function  $\rho_k(\tau)$  in (2.11b) for different low-fidelity models using different sample sizes  $n$ . All bi-fidelity solutions are of rank  $r = 10$ . In (a), minimizations occur (in ascending order of  $n$ ) at  $k = 13, 13, 7$ , and  $7$ . In (b), minimizations occur (in ascending order of  $n$ ) at  $k = 85, 85, 9$ , and  $9$ . Errors are normalized by  $\|\mathbf{H}\|$ .

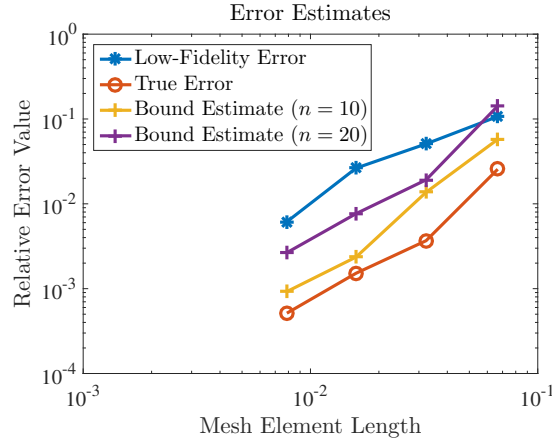


Figure 2.8: Estimate of (2.11) for various low-fidelity models and using different sample sizes  $n$ . Errors are normalized by  $\|\mathbf{H}\|$ .

are below the corresponding low-fidelity errors. In addition, the error bound estimate decays with the bi-fidelity error, as the mesh is refined.

### 2.3.2 Test Case 2: Composite Beam

For the second test case, we consider the deformation of a plane-stress, cantilever beam with composite cross section and hollow web, as shown in Figure 2.9. The Young's moduli of the three components of the cross section as well as the intensity of the applied distributed force on the beam are assumed to be uncertain, and are modeled by independent uniform random variables. Table 2.1 provides a summary of the parameters of this model along with the description of the uncertain inputs. Here, the QoI is the vertical displacement of the top cord and, in particular, its maximum occurring at the free end. To construct the bi-fidelity approximation, we use realizations of the vertical displacement over the entire top cord given by the low- and high-fidelity models.

Unlike in the previous test case where the low-fidelity model simulated the same physical problem but on a coarser mesh, here the low-fidelity model solves a simpler physical model. Specifically, the high-fidelity model is based on a finite element (FE) discretization of the beam using a triangular mesh with linear elements; see Figure 2.10. The mesh is fine enough to describe the web's hollow geometry and achieve mesh independent predictions of the vertical displacement at the free end. The FE simulations are performed using FEniCS [79]. The low-fidelity model is based on Euler-Bernoulli beam theory, where vertical cross sections are assumed to remain planes throughout the deformation and, thus, the shear deformation of the web is ignored. Additionally, we simplify the geometry of the beam by ignoring the circular holes. These simplifications, while making the calculations straightforward, lead to *inaccurate* low-fidelity predictions of the displacement (Figure 2.11), as the web experiences considerable shear deformation due to the presence of the circular holes.

Following the Euler-Bernoulli theorem, the vertical displacement of the beam  $u(x)$  is given by

$$EI \frac{d^4 u(x)}{dx^4} = -q, \quad (2.21)$$

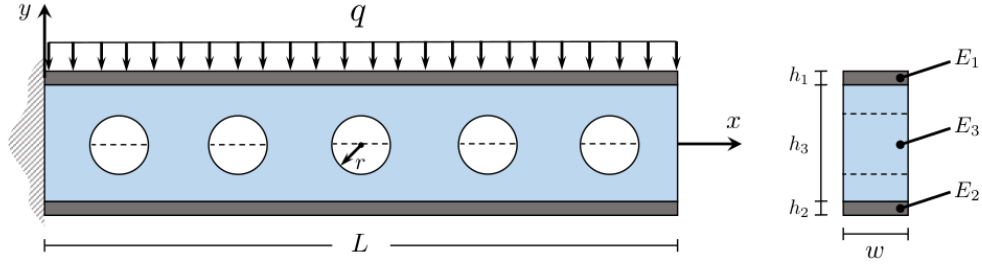


Figure 2.9: Schematic of the cantilever beam (left) and its composite cross section (right).

$L$	$h_1$	$h_2$	$h_3$	$w$	$r$	$q$	$E_1$	$E_2$	$E_3$
50	0.1	0.1	5	1	1.5	$U[9, 11]$	$U[0.9e6, 1.1e6]$	$U[0.9e6, 1.1e6]$	$U[0.9e4, 1.1e4]$

Table 2.1: Description of the parameters of the composite cantilever beam model. The center of the holes are at  $x = \{5, 15, 25, 35, 45\}$ .

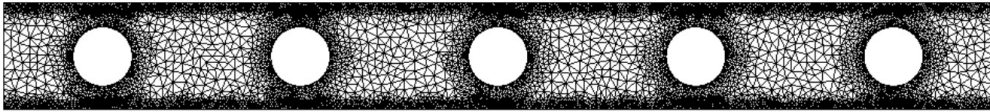


Figure 2.10: Finite element mesh used for high-fidelity simulation of the vertical displacement.

where  $E$  and  $I$  are, respectively, the Young's modulus and the moment of inertia of an *equivalent* cross section (about its centroid axis) consisting of a single material. Specifically, we let  $E = E_3$  and consider an equivalent cross section in which the width of the top and bottom sections are set to  $w_1 = (E_1/E_3)w$  and  $w_2 = (E_2/E_3)w$ , respectively, while all other dimensions remain as before; see Figure 2.9 and Table 2.1. The solution to (2.21) for the considered beam is then given by

$$u(x) = -\frac{qL^4}{24EI} \left( \left(\frac{x}{L}\right)^4 - 4\left(\frac{x}{L}\right)^3 + 6\left(\frac{x}{L}\right)^2 \right). \quad (2.22)$$

We note that the form of the function  $u(x)$  depends only weakly on the uncertain parameters; in particular the parameters only influence the magnitude of  $u(x)$  via the multiplicative coefficient  $qL^4/24EI$ . This implies that the low-fidelity solution is theoretically of rank  $r = 1$  and that only one high-fidelity realization of  $u(x)$  is needed to generate the bi-fidelity approximation. Stated differently, the bi-fidelity approximation cannot be improved by increasing the rank beyond  $r = 1$ . This has significant implications for  $\epsilon(\tau)$  in (2.9) and the corresponding interpretation of Theorem 2.2.1. Specifically, with regards to (2.11) we note that the only valid choice of  $r$  is  $r = 1$ ; therefore, both  $\sigma_{r+1}$  and  $\|\mathbf{L} - \hat{\mathbf{L}}\|$  are precisely zero so that

$$\|\mathbf{H} - \hat{\mathbf{H}}\| \leq \min_{\tau} (1 + \|\mathbf{C}_L\|) \sqrt{\epsilon(\tau)}.$$

As  $\epsilon$  is a decreasing function of  $\tau$ , we note that the minimizing value of  $\tau$  occurs in the limit as  $\tau \rightarrow \infty$ . This can be numerically observed from Figure 2.12. Even though the low-fidelity model is of a minimal rank, the high-fidelity model is itself reasonably well approximated by a representation of rank  $r = 1$ . This is seen in Figure 2.11 for two randomly selected choices of input parameters, illustrating a considerably closer agreement between the bi- and high-fidelity samples relative to the low- and high-fidelity samples. Figure 2.13 shows histograms of the relative errors of the maximum displacement predictions given by the low- and bi-fidelity models. The histograms are based on 100 random realizations of the maximum displacement. The average error achieved by the bi-fidelity model is roughly an order of magnitude smaller, illustrating that the bi-fidelity approximation is indeed considerably more accurate, in spite of using only one high-fidelity sample (as  $r = 1$ ). In

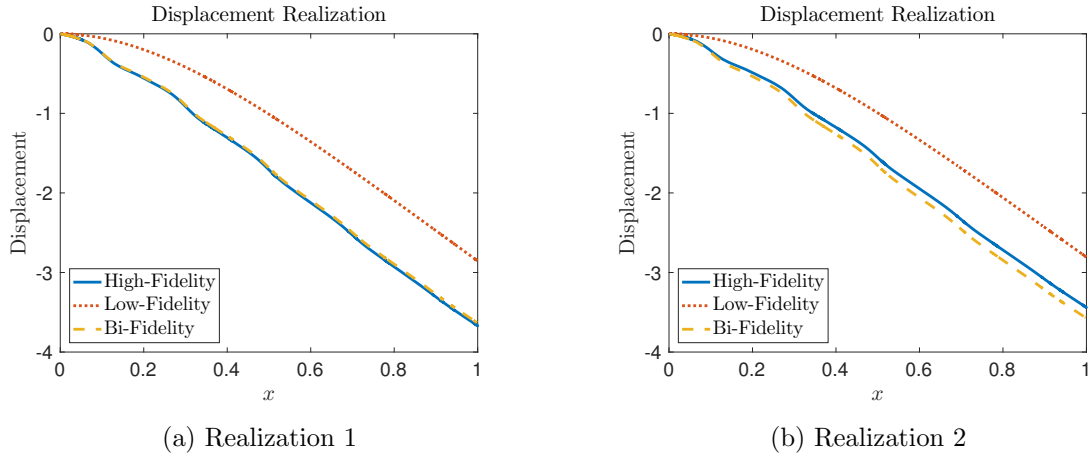


Figure 2.11: Realizations of vertical displacement for two randomly selected input parameters  $\xi$ . Shown are the low-fidelity, high-fidelity, and rank  $r = 1$  bi-fidelity estimates.

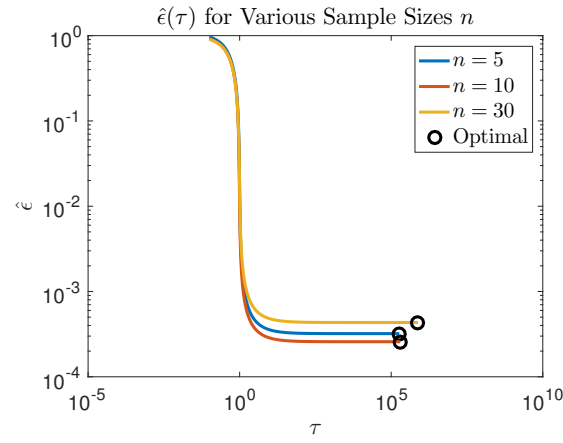


Figure 2.12:  $\hat{\epsilon}(\tau)$  from (2.16) as a function of  $\tau$  for different sample sizes  $n$ . Values which optimize (2.11) for a rank  $r = 1$  approximation are marked.



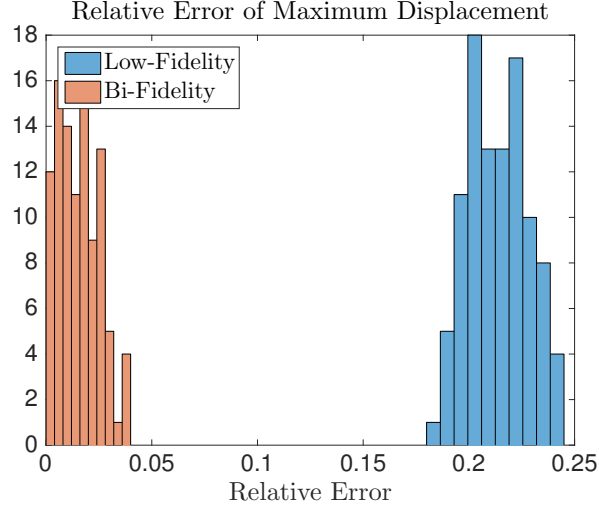


Figure 2.13: Histogram count for the relative error in the maximum displacement for low- and bi-fidelity models from 100 random realizations of the displacement.

Figure 2.14 we see that the error bound stabilizes by around 7 samples and overestimates the true error by a factor 9, which again verifies that a small number of high-fidelity model samples can reliably generate a conservative bound on the error. With regards to the efficacy of the bound (2.11) and its dependency on  $\tau$  as a function of the sample size  $n$  used to determine an optimal  $\hat{\epsilon}(\tau)$ , we see in Figure 2.15 that the bound estimate, and specifically its dependence on  $\tau$ , is robust to  $n$ , converging quickly with respect to  $n$ .

## 2.4 Conclusions

This work is concerned with the error analysis of a bi-fidelity, low-rank approximation technique for a reduced order solution of problems with stochastic or parametric inputs. The bi-fidelity method relies on the solution to a cheaper, lower-fidelity model of the intended expensive, high-fidelity model of the problem, in order to identify a reduced basis and an interpolation rule for approximating the high-fidelity solution in its inputs. A novel and practical error bound is presented for this bi-fidelity approximation that is shown to not require many high-fidelity samples beyond the rank of the approximation. This makes the bound an effective tool for practical application of this approximation when a limited computational budget is available. In particular, the

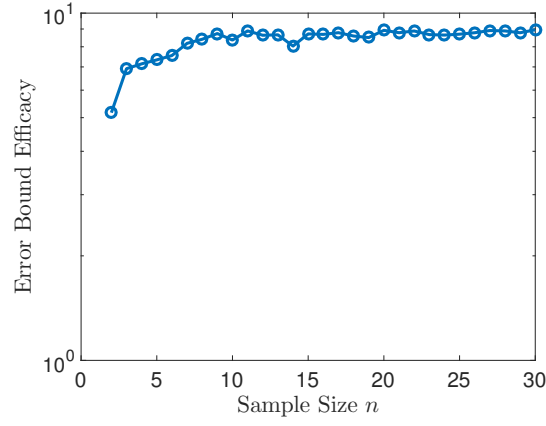


Figure 2.14: Identification of error bound efficacy, i.e., the (average) ratio between the error estimated from (2.11) and the true error, for various sample sizes  $n$ . For each  $n$  the reported error ratio is the average of 30 ratios, each computed from an independent set of  $n$  high- and low-fidelity samples via Algorithm 1.

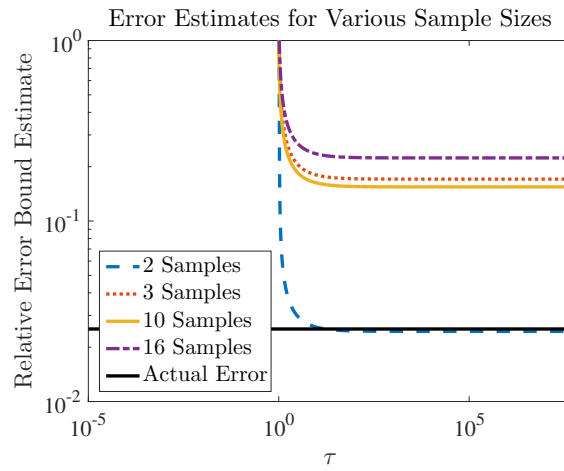


Figure 2.15: Minimum (over  $k$ ) of error function  $\rho_k(\tau)$  in (2.11b) using different sample sizes  $n$ . Minimization occurs at  $k = 1$  for all four estimates. Errors are normalized by  $\|\mathbf{H}\|$ .

error bound allows a practitioner to determine if a given pair of low- and high-fidelity models will lead to an accurate bi-fidelity approximation before investing too many computational resources. The bi-fidelity estimate and the associated error bound are by construction non-intrusive, i.e., sample-based, and thus can use legacy solvers in a black box fashion.

Some aspects of the bi-fidelity approximation, and specifically the efficacy of the derived error bound, are demonstrated on two numerical examples from fluid (non-linear) and solid (linear) mechanics. The numerical results suggest that the proposed error bound achieves a conservative estimate of the true error, while using a number of high-fidelity samples that is slightly larger than the rank of approximation. A precise estimate of the minimum number of high-fidelity samples needed to generate a conservative error bound, however, remains an open problem and requires further research.

When a low-fidelity model is not considerably cheaper than its high-fidelity counterpart, adaptive strategies for sampling the low-fidelity model are needed to identify the reduced basis and the interpolation rule with a minimal number of low-fidelity model evaluations. Additionally, deriving sharper error bounds and extending this work to include data from multiple low-fidelity models are interesting future research directions.

## Chapter 3

### A Bi-Fidelity Approximation for Uncertainty Quantification and Sensitivity Analysis of Irradiated Particle-Laden Turbulence

#### 3.1 Introduction

The ability to quantitatively characterize and reduce uncertainties, in conjunction with model verification and validation (V&V), plays a fundamental role in increasing the reliability of numerical simulations. These types of studies are commonly encompassed within the field of uncertainty quantification (UQ), which has attracted increasing attention in the modeling and simulation community. In this regard, the Predictive Science Academic Alliance Program (PSAAP) II at Stanford University [1], focuses on advancing the state-of-the-art in large-scale, predictive simulations of irradiated particle-laden turbulence relevant to concentrated solar power (CSP) systems. To this end, physics-based models are developed and the numerical predictions are validated against data acquired from an in-house experimental apparatus designed to mimic a scaled-down particle-based solar energy receiver, and for which the quantification of uncertainties is of paramount importance. A significant challenge, and the scope of this work in particular, is investigating optimal UQ strategies for this complex, multi-physics flow when many sources of uncertainty are present.

##### 3.1.1 Irradiated Particle-Laden Turbulent Flow

Turbulent flow laden with inertial particles, or droplets, in the presence of thermal radiation is encountered in a wide range of natural phenomena and industrial applications. For instance, it is well established that turbulence-driven particle inhomogeneity plays a fundamental role in

determining the rate of droplet coalescence and evaporation in ocean sprays [130] and atmospheric clouds [121]. Another example is found when studying fires, in which turbulence, soot particles, and radiation are strongly interconnected resulting in very complex physical processes [129]. From an industrial point of view, important applications include the atomization of liquid fuels in combustion chambers [75], soot formation in rocket engines [110], and more recently, volumetric particle-based solar receivers for energy harvesting [66].

Even in the simplest configuration, e.g., homogeneous isotropic turbulence, particle-laden turbulent flow is known to exhibit complex interactions between the carrier and dispersed phases in the form of preferential concentration and turbulence modulation [10]. Preferential concentration is the phenomenon by which heavy particles tend to avoid intense vorticity regions and accumulate in regions of high strain rate, while turbulence modulation refers to the alteration of fluid flow characteristics in the near-field region of particle clusters as a result of two-way coupling effects, e.g., enhanced dissipation, kinetic energy transfer, or formation of wakes and vortexes. The physical complexity is further increased by the simple addition of solid walls as turbophoresis [22], i.e., tendency of particles to migrate towards regions of decreasing turbulence levels, becomes an important mechanism for augmenting the inhomogeneity in spatial distribution of the dispersed phase by driving particle accumulation at the walls.

As described above, characterization of particle-laden turbulent flow is a difficult problem; many experimental and numerical research studies have been devoted to this objective over the past decades, e.g., [125, 132, 119]. However, the problem of interest in this work involves, in addition to particle-flow coupling, heat transfer from the particles to the fluid through radiation absorption. The practical application motivating the study of this phenomenon is the improvement of energy harvesting in volumetric particle-based solar receivers. At present, most CSP technologies use surface-based collectors to convert the incident solar radiation into thermal energy. In this type of system, the energy is transferred to the working fluid downstream of the collection point via heat exchangers, typically resulting in large conversion losses at high temperatures. By contrast, volumetric solar receivers continuously transfer the energy absorbed by particles directly to the

operating fluid as they are convected through an environment exposed to thermal radiation. This innovative technology is expected to increase the performance of CSP plants by avoiding the necessity of heat-exchanging stages, while requiring significantly high radiation-to-fluid energy transfer ratios. This requirement imposes a very complex design constraint as the physical mechanisms governing irradiated particle-laden turbulent flow are still not fully comprehended, and therefore is a topic of intense research, e.g., [139, 44].

### 3.1.2 Uncertainty Quantification for Complex, Large-Scale, High-Dimensional Systems

The complexity of constructing predictive models of CSP systems is furthered by the fact that there are many uncertainties inherent in the underlying physical processes, for instance, turbulence models, particle properties or input radiation. This high-dimensional uncertainty, in conjunction with large computational demands of simulating irradiated particle-laden turbulence, necessitates cost-efficient, non-intrusive UQ methods that accurately estimate the statistics of a given quantity of interest (QoI). As many prominent non-intrusive methods, such as stochastic collocation [88] and polynomial chaos expansions (PCEs) [48, 136, 29], are subject to the *curse of dimensionality* – the phenomenon where the computational cost grows rapidly as a function of the number of uncertainties – UQ methods within this framework must be restricted to random sampling, i.e., Monte Carlo (MC) [58], and sparse sampling methods [137]. However, for QoIs exhibiting large variability, these direct methods often become infeasible as they rely on sampling heavily from the high-fidelity (HF) model to estimate the statistics of the QoI. While the HF model accurately approximates the physical system, it is often computationally burdensome to compute, and thus much research has targeted developing stochastic model reduction techniques to improve upon this cost.

As a form of model reduction, there has been growing interest in multilevel and multi-fidelity methods, that is, methods relying on multiple models with varying levels of reliability and cost, with an aim of accurately estimating the QoI statistics in a computationally efficient manner. These

models of reduced cost and accuracy are referred to, generally, as low-fidelity (LF) models. First inspired by multigrid [15, 17], multilevel methods have evolved to not only include UQ methods exploiting coarser grid resolutions, but also multi-fidelity methods, where levels may correspond to a broader class of organization schemes, e.g., simplified physics or reduced time-stepping (the interested reader is directed to [99, 39] for a review of multi-fidelity methods). Relating to this work, interest is greatest with respect to multi-fidelity MC and sample-based interpolation methods, as the associated computational cost is not significantly impacted by an increase in the number of uncertainties.

With regards to MC, control variates [8] is a model reduction technique that introduces a second, easily simulated, and highly correlated variable as a means to reduce the variance when estimating the expected value of the QoI. This reduced variance results in fewer required HF model simulations. A specific extension of such includes the well-known multilevel Monte Carlo (MLMC) developed first in [65] and popularized by work in [49]. MLMC relies on estimating moments of the QoI via a series of coarse discretizations (temporal or spatial), and estimating the expected value via MC estimates of the expectation of QoI differences, i.e., a telescoping sum. Applications to numerical partial differential equations, as in [24, 11, 128], show the success of the method for simple mathematical systems, making it ideal for high-dimensional, large-scale problems in which there exists convergence analysis with regards to the discretization scheme. In the last decade, several other forms of control variate, in the form of multilevel and multi-fidelity control variates, have been studied which rely on LF models, many of which do not adhere strictly to the coarsening of the spatial or temporal discretization schemes, e.g., [124, 93, 131, 46, 36]. All of these model reduction methods for MC have been shown to significantly improve the computational cost in comparison to standard MC of the HF model. Interpolation methods within the multi-fidelity framework can trace their roots to the geosciences, where early work was completed in kriging and co-kriging [116]. The associated multi-fidelity methods rely on forming an autoregressive model, which forms a mapping between data obtained from the LF and HF models [72, 107, 76], treating the error between the two as a Gaussian process. Doing such requires many LF simulations, but

only a few simulations from the HF model. Similar to these methods, but absent of a Gaussian process to model the error, are bi-fidelity (BF) approximations, that rely on forming a mapping between vector-valued solutions of the LF and HF models introduced in [91, 142]. In contrast to kriging, this BF approximation forms a representative basis within the LF data, and uses this knowledge to inform the HF simulations; in addition, verification tools exist to estimate the error of this approximation [62]. An extension of this approximation in [27] attributes the success of the method to the underlying low-rank structure of the solution of interest, making it attractive for physical processes dominated by diffusion, e.g., systems where heat transfer is a dominant process.

### 3.1.3 Objectives and Organization of the Work

The system studied in this work is based on a small-scale apparatus designed to reproduce the operating conditions of volumetric particle-based solar receivers. As a consequence, many different uncertainties naturally arise when trying to numerically investigate the apparatus’s performance in terms of, for instance, thermal output and efficiency. Examples include incomplete characterization of particle-size distribution [109] and radiation properties [43], variability in radiation input and boundary conditions, and structural uncertainty inherent in the utilized models [70]. In addition to the large number of uncertainties, accurate predictions of the complex interaction of particle-laden turbulent flow with radiative heat transfer demand the utilization of expensive HF numerical simulations. As an example, the cost of a medium-scale HF calculation of this problem requires approximately 500k core-hours per sample on the Mira supercomputer (ALCF) [2]. Therefore, if brute-force UQ techniques, e.g., MC simulation with  $\mathcal{O}(10^3)$  samples, are to be performed, the total cost is of the order of 500M core-hours, motivating the need for accelerated UQ strategies. In this regard, the objective of this work is to investigate the BF approximation UQ strategy on large-scale, multi-physics applications based on the PSAAP II solar receiver. In particular, it can be shown that this BF approximation provides accurate estimates of the QoI statistics, while maintaining a reduced cost similar to that of LF models for simulations with  $\mathcal{O}(10^3)$  samples.

This chapter is organized as follows. In Section 3.2, the physical models and numerical method



utilized to simulate irradiated particle-laden turbulent flow are described. The particle-based solar receiver studied is detailed next, Section 3.3, in terms of computational setup, uncertainties, and QoIs considered. The BF approximation strategy is presented in Section 3.4, as well as a brief discussion of theoretical error results. The performance of this BF approximation, with regards to both accuracy and cost, is investigated in Section 3.5. From these results, comparisons are made between this approximation and alternative LF models. Finally, the work is concluded and future directions are proposed in Section 3.6.

## 3.2 Physics Modeling and Numerical Method

The study of volumetric particle-based solar receivers involves the interaction of particles and wall-bounded turbulent flow in a radiation environment. The equations describing this type of flow are continuity, Navier-Stokes in the low-Mach-number limit, conservation of energy assuming ideal-gas behavior, Lagrangian particle transport, and radiative heat transfer. The modeling of these three physical processes – turbulent flow, particle transport, and radiative heat transfer – and their couplings, are sequentially described in the subsections below.

### 3.2.1 Variable-Density Turbulent Flow

The volumetric particle-based solar receiver operates at atmospheric pressure conditions in which air, the carrier fluid, is assumed to follow the ideal-gas equation of state (EoS),  $P_{th} = \rho_g R_{air} T_g$ , where  $P_{th}$  is the thermodynamic pressure,  $\rho_g$  is the density,  $R_{air}$  is the specific gas constant for air, and  $T_g$  is the temperature. As indicated by the EoS, density varies with temperature. However, the Mach number of the flow  $Ma = u/c$ , with  $u$  the local flow velocity and  $c$  the speed of sound of the medium, is small ( $Ma < 0.03$ ) for the range of velocities and temperatures considered. Therefore, the low-Mach-number approximation is utilized to separate the hydrodynamic part,  $p \ll P_{th}$ , from the total pressure,  $P_{tot} = P_{th} + p$ . This decomposition results in the following

equations of fluid motion

$$\frac{\partial \rho_g}{\partial t} + \nabla \cdot (\rho_g \mathbf{u}_g) = 0, \quad (3.1)$$

$$\frac{\partial (\rho_g \mathbf{u}_g)}{\partial t} + \nabla \cdot (\rho_g \mathbf{u}_g \otimes \mathbf{u}_g) = -\nabla p + \nabla \cdot \left[ \mu_g (\nabla \mathbf{u}_g + \nabla \mathbf{u}_g^T) - \frac{2}{3} \mu_g (\nabla \cdot \mathbf{u}_g) \mathbf{I} \right] + (\rho_g - \rho_0) \mathbf{g} + \mathbf{f}_{TWC}, \quad (3.2)$$

$$\frac{\partial (\rho_g C_{v,g} T_g)}{\partial t} + \nabla \cdot (\rho_g C_{p,g} T_g \mathbf{u}_g) = \nabla \cdot (\lambda_g \nabla T_g) + S_{TWC}, \quad (3.3)$$

where  $\mathbf{u}_g$  is the gas velocity,  $\rho_0$  is an ambient reference density,  $\mathbf{I}$  is the identity matrix,  $\mathbf{g}$  is the gravitational acceleration,  $\mu_g$  and  $\lambda_g$  are the dynamic viscosity [127] and thermal conductivity [133],  $C_{v,g}$  and  $C_{p,g}$  are the isochoric and isobaric specific heat capacities, and  $\mathbf{f}_{TWC}$  and  $S_{TWC}$  are two-way coupling terms representing the effect of particles on the fluid and approximated as point sources in the form

$$\mathbf{f}_{TWC} = \sum_p m_p \frac{\mathbf{u}_p - \mathbf{v}_p}{\tau_p} \delta(\mathbf{x} - \mathbf{x}_p), \quad (3.4)$$

$$S_{TWC} = \sum_p \pi d_p^2 h (T_p - T_g) \delta(\mathbf{x} - \mathbf{x}_p), \quad (3.5)$$

where  $m_p = \rho_p \pi d_p^3 / 6$  and  $\mathbf{v}_p$  are the particle mass and velocity,  $\mathbf{u}_p$  is the gas velocity at the particle location,  $\tau_p = \rho_p d_p^2 / (18 \mu_g)$  is the particle relaxation time,  $d_p$  is the particle diameter, and  $\delta(\mathbf{x} - \mathbf{x}_p)$  is the Dirac delta function concentrated at the particle location  $\mathbf{x}_p$ . The gas-particle convection coefficient is defined as  $h = Nu \lambda_g / d_p$  with  $Nu$  the particle Nusselt number, and which takes into account the fact that the Biot number is  $Bi = h d_p / \lambda_p \ll 1$  in this problem, and therefore particles are assumed to be isothermal with temperature  $T_p$ .

### 3.2.2 Lagrangian Particle Transport

The carrier fluid is transparent to the incident radiation. Hence, micron-sized nickel particles, i.e., the dispersed phase, are seeded into the gas to generate a non-transparent gas-particle mixture that absorbs and transfers the incident radiation from the particles to the gas phase. The diameters of the particles are several orders of magnitude smaller than the smallest significant (Kolmogorov)

turbulent scale  $\tau_\eta$ , and the density ratio between particles and gas is  $\rho_p/\rho_g \gg 1$ . As a result, particles are modeled following a Lagrangian point-particle approach with Stokes' drag as the most important force [89]. Their description in terms of position, velocity and temperature is given by

$$\frac{d\mathbf{x}_p}{dt} = \mathbf{v}_p, \quad (3.6)$$

$$\frac{d\mathbf{v}_p}{dt} = \frac{\mathbf{u}_p - \mathbf{v}_p}{\tau_p} + \mathbf{g}, \quad (3.7)$$

$$\frac{d(m_p C_{v,p} T_p)}{dt} = \frac{\pi d_p^2 (1 - \omega)}{4} \int_{4\pi} \left( I - \frac{\sigma T_p^4}{\pi} \right) d\Omega - \pi d_p^2 h (T_p - T_g), \quad (3.8)$$

where  $C_{v,p}$  is the particle specific isochoric heat capacity,  $\omega = Q_s/(Q_a + Q_s)$  is the scattering albedo with  $Q_a$  and  $Q_s$  the absorption and scattering efficiencies, respectively,  $I$  is the radiation intensity,  $\sigma$  is the Stefan-Boltzmann constant, and  $d\Omega = \sin\theta d\theta d\phi$  is the differential solid angle. In the conservation equation for particle temperature, (3.8), the first term on the right-hand-side accounts for the amount of radiation absorbed by a particle, while the second term represents the heat transferred to its surrounding fluid.

In the point-particle approximation, particle-wall and particle-particle interactions are typically described by one-dimensional collision models based on the balance of total momentum and energy. In the case of collisions involving two objects,  $A$  and  $B$ , the velocities after impact,  $v_A$  and  $v_B$ , are given by

$$v_A = [m_A u_A + m_B u_B + m_B C_R (u_B - u_A)] / (m_A + m_B), \quad (3.9)$$

$$v_B = [m_A u_A + m_B u_B + m_A C_R (u_A - u_B)] / (m_A + m_B), \quad (3.10)$$

where  $u_A$  and  $u_B$  are the velocities of the objects before impact,  $m_A$  and  $m_B$  are the mass of the objects, and  $0 \leq C_R \leq 1$  is the restitution coefficient. The limits of  $C_R$  correspond to the cases in which the objects coalesce at impact (0, perfectly inelastic collision) and rebound with the same relative speed as before impact (1, perfectly elastic collision); intermediate values represent inelastic collisions in which kinetic energy is dissipated. The above equations simplify to  $v_A = -C_R u_A$  and  $v_B = 0$  when object  $B$  is a static wall.

### 3.2.3 Radiative Heat Transfer

In the problem under consideration, the flow and particles timescales are orders of magnitude larger than the radiation timescale, which is related to the speed of light. As a consequence, it can be assumed that the radiation field changes instantaneously with respect to temperature and particle distributions; i.e., radiation field is quasi-steady. Under this assumption, and considering that air is transparent at all wavelengths and that absorption and scattering are determined solely by the presence of particles and solid boundaries, the radiative heat transfer equation becomes

$$\hat{\mathbf{s}} \cdot \nabla I = -\sigma_e I + \sigma_a \frac{\sigma T_p^4}{\pi} + \frac{\sigma_s}{4\pi} \int_{4\pi} I \Phi d\Omega, \quad (3.11)$$

where  $\hat{\mathbf{s}}$  is the direction vector,  $\sigma_e = \sigma_a + \sigma_s$  is the total extinction coefficient with  $\sigma_a$  and  $\sigma_s$  the absorption and scattering coefficients, respectively, and  $\Phi$  is the scattering phase function that describes the directional distribution of scattered radiation. The total extinction coefficient can also be defined in terms of absorption and scattering efficiencies as  $\sigma_e = (Q_a + Q_s) \pi d_p^2 n_p / 4$  with  $n_p$  the local number density of particles. Moreover, assuming gray radiation  $Q_a + Q_s \approx 1$ , which leads to  $\omega \approx Q_s$ , and as a result  $\sigma_a \approx Q_a \pi d_p^2 n_p / 4$  and  $\sigma_s \approx Q_s \pi d_p^2 n_p / 4$ .

### 3.2.4 Numerical Method

The equations of fluid motion, (3.1)-(3.3), are solved following an Eulerian finite-volume discretization implemented in an in-house solver that is second-order accurate in space and suitable to non-uniform meshes. A fourth-order Runge-Kutta scheme is used for integrating the equations in time, together with a fractional-step method for imposing conservation of mass [35]. Integration in time of the Lagrangian position, velocity, and temperature of particles, (3.6)-(3.8), is fully coupled with the advancement of the flow equations to ensure fourth-order accuracy. The transfer of radiative heat, (3.11), is calculated by means of an in-house discrete ordinates method (DOM) interfaced to the flow solver by means of an Eulerian representation of the particles distribution.

### 3.3 Description of the Particle-Based Solar Receiver

#### 3.3.1 Computational Setup and Physical Parameters

Numerical simulations of the volumetric particle-based solar receiver are performed on the computational setup depicted in Figure 3.1. Two square duct domains, with dimensions  $1.7L \times W \times W$  ( $L = 0.16$  m,  $W = 0.04$  m) in the streamwise ( $x$ -axis) and wall-normal directions ( $y$ - and  $z$ -axis), are utilized to mimic the development and radiated sections of the experimental apparatus. The development section (left domain) is an isothermal,  $T_0$ , periodic particle-laden turbulent flow generator that provides inlet conditions for the inflow-outflow radiated section (right domain). The solid boundaries of the development section ( $y$ - and  $z$ -sides) are considered smooth, no-slip, adiabatic walls. Regarding the radiated section, the same boundary conditions are imposed except for the radiated region in which the  $y$ - and  $z$ -boundaries are modeled as non-adiabatic walls accounting for heat fluxes due to the radiation energy absorbed by the glass windows.

The bulk Reynolds number of the gas phase at the development section is  $Re_b = \rho_g u_b L / \mu_g = 20,000$ , with  $u_b$  the gas bulk velocity. The particle-size distribution is approximated by 5 different classes with Kolmogorov Stokes numbers in the range  $5 < St_\eta = \tau_p / \tau_\eta < 20$  and with a total mass loading ratio (MLR) of  $n_p m_p / \rho_g \approx 20\%$ . Detailed values of the development section flow conditions and material properties are listed in Table 3.1. The gas-particle mixture, as depicted in Figure 3.2, is volumetrically irradiated through an  $L \times W$  glass window starting at  $\Delta x = 0.1L$  from the beginning of the radiated section. The radiation source consists of an array of diodes mounted on a vertical support placed  $\Delta y = 2.875W$  from the radiated window and aligned with the streamwise direction of the flow. The diodes generate a total power of  $P \approx 1$  kW approximately uniform within a  $18^\circ$  cone angle.

#### 3.3.2 Uncertainties and Quantities of Interest

The uncertainty quantification study considers 14 stochastic variables to describe experiment and model-form uncertainties, as shown in Table 3.2. These correspond to incertitude in particle

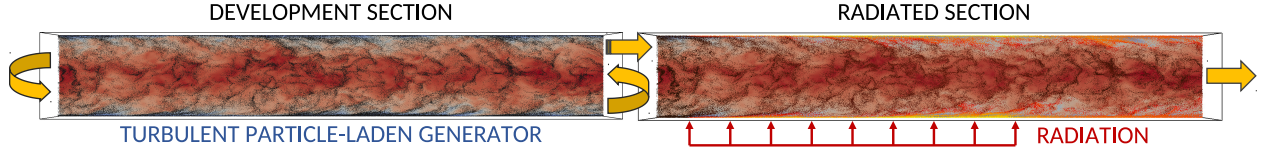


Figure 3.1: Computational setup of the PSAAP II volumetric particle-based solar energy receiver. An isothermal periodic section (left domain) is utilized to generate fully developed particle-laden turbulent flow, which is used as inflow conditions for the second section (right domain) where the gas-particle mixture is irradiated perpendicularly to the flow direction from one the sides.

Parameter	Value	Parameter	Value
$u_b$	8 m/s	$T_0$	300 K
$P_{th}$	101325 Pa	$R_{air}$	287 J/(kg·K)
$C_{p,g}$	1012 J/(kg·K)	$C_{v,g}$	723 J/(kg·K)
$\rho_p$	8900 kg/m <sup>3</sup>	$C_{v,p}$	450 J/(kg·K)
$d_p$	8.4, 9.8, 11.2, 12.2, 14.6 $\mu\text{m}$	$\mathbf{g}$	(9.81, 0, 0) m/s <sup>2</sup>

Table 3.1: Flow conditions at development section and physical properties.

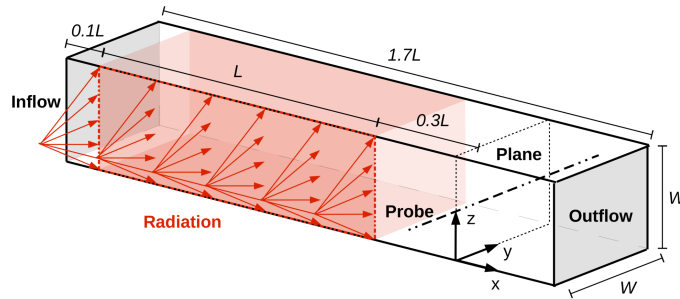


Figure 3.2: Detailed schematic of the radiated test section. The domain is  $1.7L$  long and  $W$  width, with the radiated region extending  $\Delta x = L$  and  $\Delta z = W$  starting at  $\Delta x = 0.1L$  from the beginning of the section, and with a probe perpendicular to the flow and located  $\Delta x = 0.3L$  downstream.

Variable	Interval	Variable	Interval
1. Prt. rest. coeff. 1	[0.0 : 0.6]	8. Mass load. ratio	[18 : 22]%
2. Prt. rest. coeff. 2	[0.1 : 0.7]	9. Prt. abs. eff.	[0.37 : 0.41]
3. Prt. rest. coeff. 3	[0.2 : 0.8]	10. Prt. scatt. eff.	[0.69 : 0.76]
4. Prt. rest. coeff. 4	[0.3 : 0.9]	11. Radiation	[1.8 : 2.0] MW/m <sup>2</sup>
5. Prt. rest. coeff. 5	[0.4 : 1.0]	12. Radiated wall	[1.6 : 6.4] kW/m <sup>2</sup>
6. Stokes' drag corr.	[1.0 : 1.5]	13. Opposite wall	[1.2 : 4.7] kW/m <sup>2</sup>
7. Prt. Nusselt num.	[1.5 : 2.5]	14. Side $x$ - $y$ walls	[0.1 : 0.2] kW/m <sup>2</sup>

Table 3.2: List of random inputs with the corresponding ranges. All inputs are assumed to be uniformly distributed.

restitution coefficient for the different classes (1 – 5), correction to Stokes' drag law (6), particle Nusselt number (7), mass loading ratio (8), particle absorption and scattering efficiencies (9 – 10), incident radiation flux (11), and heat fluxes from the walls to the fluid (12 – 14).

The intervals of the stochastic variables listed in Table 3.2 have been carefully characterized on the basis of information provided by the team responsible for conducting the experiments, and by taking into consideration results and conclusions extracted from published studies. The intervals of the particle restitution coefficients follow the trend observed in experimental investigations by Yang & Hunt [138] in which  $C_R$  increases with Stokes number. The expression for Stokes' drag force correction and its coefficient interval is based on the theoretical work by Brenner [16]. The particle Nusselt number range is extracted from numerical experiments of heated particles. The intervals for particle absorption and scattering efficiencies are obtained from Mie scattering theory and take into account sensitivity to shape deformation as investigated by Farbar et al. [38]. Intervals for mass loading ratio, incident radiation, and heat fluxes from the walls to the fluid are characterized based on comparisons between preliminary numerical simulations and experimental results.

Time-averaged three-dimensional (3-D) solutions of the numerical simulations are saved in binary files from which first- and second-order statistics of different QoIs can be analyzed. For example, gas velocity and density distributions, mean gas temperature and fluctuations, transmitted light, number density, velocity and temperature of particles, etc. However, in this work, the performance of the BF approximation is focused on thermal QoIs at the probe location. As detailed

in Figure 3.2, the probe is located  $\Delta x = 0.3L$  downstream from the radiated perimeter, and is perpendicular to the flow direction along the  $y$ -axis at  $z = W/2$ . Of particular interest, as this quantity is available from the experiment, is the time-averaged  $y$ -axis profile of gas temperature.

### 3.3.3 Simulation Strategy and Ensemble Calculations

In the experimental apparatus, fully developed conditions are achieved using a long duct with an aspect ratio of order one hundred. To reduce the computational cost of simulating a long duct with inflow-outflow boundary conditions, the computational setup is divided in two domains as described in Section 3.3.1. First, randomly distributed particles are seeded into the development section with an initial fully developed turbulent velocity field. This system is then evolved in time for 20 flow through times (FTTs), defined as  $FTT = L/u_b$ , to achieve fully developed turbulent particle-laden flow conditions as in the experiments. After 20 FTTs, the instantaneous Eulerian and Lagrangian solutions are copied into the second domain, radiative illumination is activated in the radiated region, and the total system (two domains) is integrated in time for 5 additional FTTs intended to flush the thermal transient (1 FTT) and collect statistics (4 FTTs). This procedure is repeated independently for each realization.

Quantification of uncertainties in complex systems by means of non-intrusive sampling methods, like the solar receiver UQ study of this work, typically requires on the order of a thousand model evaluations for different input values. Leadership computing facilities prefer a small number of large calculations rather than a large number of small jobs. In this regard, the UQ analysis has been performed by grouping multiple realizations into few large ensemble jobs. This methodology is highly prioritized by the job resource manager resulting in reduced turn-around times by minimizing wait-time and providing larger run-time allowance.

### 3.3.4 Description of the High- and Low-Fidelity Models

Three fidelity levels have been designed to perform the UQ study: one HF model and two LF representations, denoted LF1 and LF2. The HF level corresponds to a point-particle direct numer-



ical simulation (PP-DNS) with sufficient resolution ( $\approx 55\text{M}$  cells/section) to capture all significant (integral to Kolmogorov) turbulent scales, while approximating the particles as Lagrangian points ( $\approx 15\text{M}$  particles/section) with nonzero mass. The flow grid is uniform in the streamwise direction with spacings in wall units equal to  $\Delta x^+ \approx 12$ , while stretched in the wall-normal directions with the first grid point at  $y^+, z^+ \approx 0.5$  and with resolutions in the range  $0.5 < \Delta y^+, \Delta z^+ < 6$ . The radiative heat transfer equation is solved on a DOM mesh of  $270 \times 160 \times 160$  gridpoints ( $\approx 7\text{M}$  cells) with 350 quadrature points (discrete angles).

Based on the HF model, two LF levels have been constructed by carefully coarsening the Eulerian and Lagrangian resolutions (allowing, in addition, for larger time steps), resulting in the LF1 and LF2 representations that are  $\approx 170\times$  and  $\approx 1300\times$  cheaper per sample than HF, respectively. A depiction of the different fidelity levels for the turbulent gas phase is shown in Figure 3.3. To explore the acceleration strategies, HF, LF1, and LF2 samples derived from identical input realizations have been computed using resources from Titan (OLCF) [3] and Mira (ALCF) [2] supercomputers.

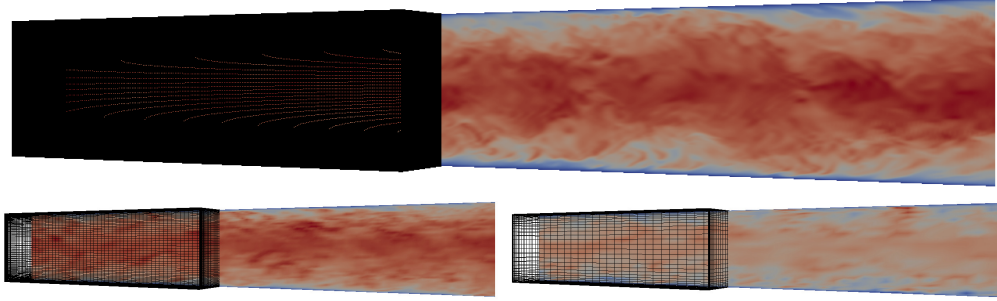


Figure 3.3: Turbulent gas phase Eulerian resolution of the fidelity levels designed; gas velocity as background. HF  $540 \times 320 \times 320$  gridpoints (top), LF1  $108 \times 64 \times 64$  gridpoints (bottom left), LF2  $54 \times 32 \times 32$  gridpoints (bottom right).

### 3.4 Bi-fidelity Approximation Strategy

#### 3.4.1 Formation of Bi-fidelity Approximation

The BF approximation described in this work follows from the description in [27], which exploits the low-rank structure of LF and HF QoI solutions by forming a low-rank approximate mapping between the two. This mapping treats the  $d$  uncertainties ( $d = 14$  in this study, as described in Table 3.2) as components of a random parameter  $\boldsymbol{\xi} = (\xi_1, \dots, \xi_d)$  with an assumed joint probability density function (PDF) (in this study, a uniform joint PDF). While the QoIs in this work are scalar-valued, this approximation relies on vector-valued quantities  $\mathbf{u}$  that describe the QoI. For example, the elements of  $\mathbf{u}$  may be numerical estimates of time-averaged  $y$ -axis profiles of the gas temperature, and the QoI  $Q$  is the spatially-averaged gas temperature along this profile. The reader is referred to Figure 3.4 for an illustrated description of the steps required to form this BF approximation. The construction of the BF approximation is completed in four main steps. In

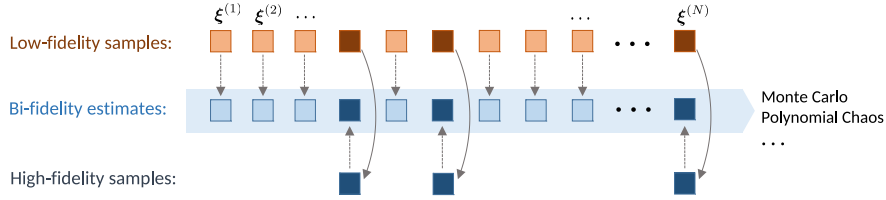


Figure 3.4: Schematic for the formation of the BF approximation

the first step,  $N$  i.i.d. inputs  $\boldsymbol{\xi}^{(i)}$  are selected according to their joint PDF and used to form  $N$  LF realization outputs  $\mathbf{u}_L^{(i)}$  (see Figure 3.4, orange boxes, top row). In matrix form, the  $m \times N$  LF data matrix, whose columns are formed by the realizations of  $\mathbf{u}_L^{(i)}$ , is defined as

$$\mathbf{U}_L := \begin{bmatrix} \mathbf{u}_L^{(1)} & \mathbf{u}_L^{(2)} & \dots & \mathbf{u}_L^{(N)} \end{bmatrix}. \quad (3.12)$$

Typically  $N$  is large, as obtaining many LF samples is computationally feasible, and the value of  $m$  corresponds to the spatial degrees of freedom of the numerical solver. In the second step, a rank  $r \ll N$  matrix interpolative decomposition (MID) [54, 23, 84] is applied to  $\mathbf{U}_L$ . The rank  $r$

approximation is given by

$$\hat{U}_L := \mathbf{Q} \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \end{bmatrix} \mathbf{P},$$

where  $\mathbf{Q} \in \mathbb{R}^{m \times m}$  is an orthogonal matrix,  $\mathbf{R}_{11} \in \mathbb{R}^{m \times r}$  is an upper triangular matrix,  $\mathbf{R}_{12} \in \mathbb{R}^{m \times N-r}$  is a dense matrix, and  $\mathbf{P} \in \mathbb{R}^{N \times N}$  is a permutation matrix. The product  $\mathbf{Q}\mathbf{R}_{11}$  forms the  $r$  columns of  $\mathbf{U}_L$  that act as a representative basis for the LF matrix. Setting  $\mathbf{U}_L^c := \mathbf{Q}\mathbf{R}_{11}$  to be the  $m \times r$  *column skeleton* of  $\hat{U}_L$ , it follows that

$$\mathbf{U}_L^c = \begin{bmatrix} \mathbf{u}_L^{(i_1)} & \mathbf{u}_L^{(i_2)} & \cdots & \mathbf{u}_L^{(i_r)} \end{bmatrix}, \quad (3.13)$$

where the indices  $\{i_1, \dots, i_r\}$  indicate the LF vectors that form the column skeleton, as well as the  $r$  inputs  $\{\boldsymbol{\xi}^{(i_1)}, \dots, \boldsymbol{\xi}^{(i_r)}\}$  that form these realizations (see Figure 3.4, dark orange boxes, top row). Then, setting  $\mathbf{C}_L = [\mathbf{I}|\mathbf{R}_{11}^+\mathbf{R}_{12}]\mathbf{P}$  to be the LF coefficient matrix, the rank  $r$  MID approximation of  $\mathbf{U}_L$  is given as

$$\hat{U}_L := \mathbf{U}_L^c \mathbf{C}_L. \quad (3.14)$$

In the third step, the inputs that provide the LF realizations of  $\mathbf{U}_L^c$  in (3.13) are applied to the HF model (see Figure 3.4, dark blue boxes, bottom row). This results in the associated HF column skeleton

$$\mathbf{U}_H^c = \begin{bmatrix} \mathbf{u}_H^{(i_1)} & \mathbf{u}_H^{(i_2)} & \cdots & \mathbf{u}_H^{(i_r)} \end{bmatrix}. \quad (3.15)$$

Note that  $\mathbf{U}_H^c$  is  $M \times r$ , where  $M \geq m$  is the value corresponds to the spatial degrees of freedom. For HF models with a finer spatial mesh resolution than the LF model it often follows that  $M > m$ ; however, as is the case with this work, data values may be extracted at equivalent coordinates, resulting in  $M = m$ . In the final step the BF approximation is formed by taking the product of the HF column skeleton in (3.15) and the LF coefficient matrix

$$\hat{U}_H := \mathbf{U}_H^c \mathbf{C}_L, \quad (3.16)$$

where a column of  $\hat{U}_H$ , denoted  $\hat{\mathbf{u}}_H^{(i)}$ , is the BF approximation to  $\mathbf{u}_H^{(i)}$  (see Figure 3.4, blue boxes, middle row). Once formed, approximate QoI realizations are calculated directly from each sample of  $\hat{\mathbf{u}}_H^{(i)}$ .

With regards to computational cost, the BF approximation requires  $N$  LF simulations, performing rank revealing QR with  $\mathcal{O}(rmN)$  floating point operations for basis identification and formation of the coefficient matrix, and only  $r$  HF simulations to form the columns of  $\mathbf{U}_H^c$ . Typically, obtaining HF solutions is a bottleneck, and thus limiting the number of HF simulations to small  $r$  is of great value and a fundamental component of this approximation.

**Remark 3.4.1.** *While the optimal rank for this BF approximation is not known a priori, the rank of the LF data matrix  $\mathbf{U}_L$  may indicate a good selection range. Theoretical results of [54, 23, 84] show the error of the rank  $r$  LF approximation in (3.14) is bounded by a scaling of the  $r+1$  largest singular value of  $\mathbf{U}_L$ . This suggests that the decay of singular values of  $\mathbf{U}_L$ , specifically, where a significant drop occurs, to be a worthwhile consideration for the value of  $r$ .*

### 3.4.2 Theoretical Error Estimation of Bi-Fidelity Approximation

For systems exhibiting a low-rank structure, this BF approximation has proven to be successful [91, 142, 27, 122, 62]; however, theoretical verification tools have only recently been developed in [62]. This section briefly presents the theoretical results of [62] that provide an error bound on the spectral norm of  $\mathbf{U}_H - \hat{\mathbf{U}}_H$ .

Define  $\mathbf{U}_L = \mathbf{U}\Sigma\mathbf{V}^T$  to be the singular value decomposition of  $\mathbf{U}_L$ , and  $\mathbf{V}_k$  be the  $k$  columns of  $\mathbf{V}$  corresponding to the  $k$  largest singular values. Let  $\mathbf{P}_{\mathbf{V}_k} = \mathbf{V}_k\mathbf{V}_k^T$  be the orthogonal projection onto the space spanned by  $\mathbf{V}_k$ . The matrix  $\mathbf{T}_k$ , defined in [62], is the operator that maps the first  $k$  singular vectors of  $\mathbf{U}_L$  to  $\mathbf{U}_H$ , given as

$$\mathbf{T}_k = \mathbf{U}_H \mathbf{P}_{\mathbf{V}_k} \mathbf{U}_L^\dagger, \quad (3.17)$$

where  $^\dagger$  indicates the pseudoinverse. Consider the mapping between the LF and HF matrices, given by  $\mathbf{U}_H = \mathbf{T}_k \mathbf{U}_L + \mathbf{E}$ , where  $\mathbf{E}$  corresponds to the error of the approximate mapping and is defined to be the modes of  $\mathbf{U}_H$  that are in the orthogonal complement of  $\mathbf{P}_{\mathbf{V}_k}$ . To account for the error of the mapping, the set of points  $(\tau, \epsilon(\tau))$  consider the proximity of the HF and LF Gramians

$$\epsilon(\tau) = \lambda_{\max}(\mathbf{U}_H^T \mathbf{U}_H - \tau \mathbf{U}_L^T \mathbf{U}_L). \quad (3.18)$$

While this relationship is not derived directly from the assumption  $\mathbf{U}_L = \mathbf{T}_k \mathbf{U}_L + \mathbf{E}$ , the parameters  $\tau$  and  $\epsilon(\tau)$  capture the behavior of this mapping, and provide an error bound on the norm of  $\mathbf{U}_H - \mathbf{T}_k \mathbf{U}_L$  and the norm of the operator  $\mathbf{T}_k$ .

The following theorem provides an error bound for the BF approximation in (3.16) with regards to the HF data.

**Theorem 3.4.1** (Theorem 1 of [62]). *Let  $\hat{\mathbf{U}}_H$  be the rank  $r$  BF approximation, as in (3.16), to the HF data matrix  $\mathbf{U}_H$ . Let  $\hat{\mathbf{U}}_L$  be the rank  $r$  MID approximation, given in (3.14), to the LF data matrix  $\mathbf{U}_L$ , where  $\mathbf{C}_L$  is the corresponding coefficient matrix. For  $\epsilon(\tau)$ , as defined in (3.18), and  $\|\cdot\|$  the spectral norm, it follows that the BF error may be bounded as*

$$\|\mathbf{U}_H - \hat{\mathbf{U}}_H\| \leq \min_{\substack{k < \text{rank}(\mathbf{U}_L) \\ \tau \geq 0}} \left( (1 + \|\mathbf{C}_L\|) \sqrt{\tau \sigma_{k+1}^2 + \epsilon(\tau)} + \|\mathbf{U}_L - \hat{\mathbf{U}}_L\| \sqrt{\tau + \epsilon(\tau) \sigma_k^{-2}} \right), \quad (3.19)$$

where  $\sigma_k$  and  $\sigma_{k+1}$  are the  $k$  and  $k+1$  largest singular values of  $\mathbf{U}_L$ .

As a brief outline to the proof (see [62] for complete details), it follows that the error may be bounded as

$$\begin{aligned} \|\mathbf{U}_H - \hat{\mathbf{U}}_H\| &\leq \|\mathbf{U}_H - \mathbf{T}_k \mathbf{U}_L\| + \|\mathbf{T}_k \mathbf{U}_L - \mathbf{T}_k \hat{\mathbf{U}}_L\| + \|\mathbf{T}_k \hat{\mathbf{U}}_L - \hat{\mathbf{U}}_H\| \\ &\leq (1 + \|\mathbf{C}_L\|) \|\mathbf{U}_H - \mathbf{T}_k \mathbf{U}_L\| + \|\mathbf{U}_L - \hat{\mathbf{U}}_L\| \|\mathbf{T}_k\|. \end{aligned} \quad (3.20)$$

Based on the definition of  $\mathbf{T}_k$  in (3.17) and  $\epsilon(\tau)$  in (3.18), it can be shown that  $\|\mathbf{U}_H - \mathbf{T}_k \mathbf{U}_L\|$  and  $\|\mathbf{T}_k\|$  are bounded by the square root terms in (3.19).

The true values of  $(\tau, \epsilon(\tau))$  require access to all  $N$  columns of  $\mathbf{U}_H$  and  $\mathbf{U}_L$ , which may be infeasible. As an alternative, estimates may be calculated by using a small subset of  $R$  HF and LF samples via

$$\hat{\epsilon}(\tau) = \frac{N}{R} \lambda_{\max}((\mathbf{U}_H^R)^T \mathbf{U}_H^R - \tau (\mathbf{U}_L^R)^T \mathbf{U}_L^R), \quad (3.21)$$

where the superscript  $R$  indicates the number of columns. To estimate the remainder of the bound in (3.19), MID is applied to the LF data matrix to obtain values for  $\|\mathbf{U}_L - \hat{\mathbf{U}}_L\|$  and  $\|\mathbf{C}_L\|$ . Combining these estimates and minimizing over identified values of  $(\tau, \hat{\epsilon}(\tau))$  and the singular values

of  $\mathbf{U}_L$  results in an approximate bound of (3.19). For an effective bound, normalization of  $\mathbf{U}_L$  may be required to estimate values of  $(\tau, \hat{\epsilon}(\tau))$ , and subsequently (3.21). In this work normalization of  $\mathbf{U}_H$  and  $\mathbf{U}_L$  is completed with respect to the Frobenius norms, i.e.,  $\mathbf{U}_H := \mathbf{U}_H \|\mathbf{U}_H\|_F^{-1}$  and  $\mathbf{U}_L := \mathbf{U}_L \|\mathbf{U}_L\|_F^{-1}$ .

### 3.4.3 Bi-Fidelity Approximation for Estimation of QoI Statistics

Finally, the application of this approximation in practice is discussed. When the rank  $r$  BF approximation informs the HF data well, methods such as MC, control variates, sparse grid collocation, or sparse PCE may be employed on the BF data to estimate the moments and PDF of the QoI, and perform global sensitivity analysis. For high-dimensional systems, strategies for estimating the statistics of interest are restricted to MC and control variates. If the uncertainty of the system is not too high, e.g., obtaining a large number of BF samples is feasible (as is the case in this work), a sparse PCE surrogate may be formed, or sparse grid collocation may be applied. On the other hand, if the rank  $r$  BF approximation has a large bias, but is well correlated with the HF data, the BF approximation may serve as a control variate to MC in a single-level or a multilevel setting as in [36]. Due to the low bias observed in this work, the focus is placed on the former implementation strategy.

As obtaining many BF simulations is dependent on the computational cost of the LF model, and thus feasible, a second order sparse PCE is employed [29, 60] to estimate the statistics of interest of the QoI, including the mean, variance, and PDF, as well as perform global sensitivity analysis via Sobol' indices. A PCE of the QoI, which provides a mapping between the random parameter  $\boldsymbol{\xi}$  and the QoI, is formed by projecting the QoI onto a polynomial basis [136]. When forming a sparse PCE, the polynomial coefficients are approximated by fitting the data to the polynomials, such that sparsity is enforced. Once the PCE coefficients have been approximated, the mean is identified from the first coefficient and the variance from the sum of the remaining squared coefficients. The complete sparse PCE surrogate is used to estimate the QoI PDF, in the form of a histogram. This is done by applying many additional samples of  $\boldsymbol{\xi}$  to the PCE surrogate,

which is done at negligible cost, and then using the QoI estimates to form a histogram.

Global sensitivity analysis may be performed via Sobol' indices, where the indices are calculated from the approximate PCE coefficients [126, 56]. This variance decomposition provides a method to quantitatively describe the importance of input parameters, by calculating each parameter's contribution to the total variance of the output QoI.

### 3.5 Numerical Results of Bi-Fidelity Approximation

To investigate the performance of the BF approximation, 256 LF2, 128 LF1, and 26 HF simulations were performed, such that the 26 HF simulations correspond to the first 26 simulations of LF1 and LF2, and the 128 LF1 simulations correspond to the first 128 of the LF2 simulations. From the two LF models, two BF approximations are formed: bi-fidelity 1 (BF1) approximation and bi-fidelity 2 (BF2) approximation. The BF1 approximation is formed from  $N = 128$  LF1 samples and  $r$  HF samples, and the BF2 approximation is formed from  $N = 256$  LF2 samples and  $r$  HF samples. The number of HF simulations is left as  $r$ , as the selection of this value will be discussed in the following results.

The motivation of this approximation is to form a reduced model that accurately describes the HF data, and the goal of these results is to investigate whether or not there is an improvement over the performance of the the LF models. For these tests two primary time-averaged thermal QoIs are considered: (i) total normalized heat flux through the plane at the probe location ( $\Delta x = 0.3L$  downstream from the radiated perimeter), and (ii)  $\Delta T/T_0 = (T - T_0)/T_0$  values along the profile at the probe location ( $\Delta x = 0.3L$  downstream from the radiated perimeter at  $z = 0.5W$ ), where focus is placed on spatially-averaged  $\Delta T/T_0$ , and point estimates at  $y/W = 0.5$ ,  $y/W = 0.1$ , and  $y/W = 0.05$ , as these are quantities available from experiments.

For each QoI, five primary tasks are considered: (i) BF rank identification as to best optimize accuracy and computational cost of the approximation, (ii) BF error bound estimation, to verify the accuracy of the approximation for a fixed rank, (iii) QoI approximation via available data, (iv) estimation of statistics via sparse PCE of LF and BF models, and (v) cost analysis to compare

approximate core-hours needed to obtain converged simulations from each model.

Errors are estimated via relative spectral error for vector-valued quantities, and relative  $\ell_2$  error for scalar quantities. For a given matrix  $\mathbf{U}$  (of LF or BF data), define the relative spectral error to be

$$\text{relative spectral error} = \frac{\|\mathbf{U}_H - \mathbf{U}\|}{\|\mathbf{U}_H\|}, \quad (3.22)$$

where  $\mathbf{U}_H$  is the matrix of HF data. Define the relative  $\ell_2$  error to be

$$\text{relative } \ell_2 \text{ error} = \frac{\sqrt{\sum_{i=1}^N (Q_H^{(i)} - Q^{(i)})^2}}{\sqrt{\sum_{i=1}^N (Q_H^{(i)})^2}}, \quad (3.23)$$

where  $Q^{(i)}$  is a LF or BF simulated QoI and  $Q_H^{(i)}$  is the corresponding HF simulated QoI.

### 3.5.1 QoI #1: Heat Flux Through the $\Delta x = 0.3L$ Plane

To estimate the statistics of the heat flux QoI the BF approximation is formed from realizations  $\mathbf{u}$  of the heat flux values over the entire  $\Delta x = 0.3L$  plane. For ease of comparison, these values are extracted as a  $500 \times 500$  uniform grid of points for both HF and LF simulations. The scalar-valued QoI estimates  $Q$  are taken to be the total heat flux, i.e., the sum, as determined from the elements of  $\mathbf{u}$ , normalized by the mean QoI estimated from the HF data.

When forming the BF approximation, the first task is to identify the approximation rank. This is critical as it will dictate both the computational cost and accuracy of the approximation. As a reduced cost relies on minimal HF samples, the selected rank must be limited such that it captures the important modes, and no more. On the other hand, improved accuracy may be observed for higher ranks. To optimize both cost and accuracy, the minimum rank that allows for the capture of important modes of the LF data must be identified. To aid in rank selection, consider the results of Figure 3.5. Figure 3.5 (a) provides the decay of singular values of LF and HF matrices. The magnitudes of these singular values decay most rapidly within the first six indices, indicating that the first six modes capture most of the information of the LF and HF data. Figure



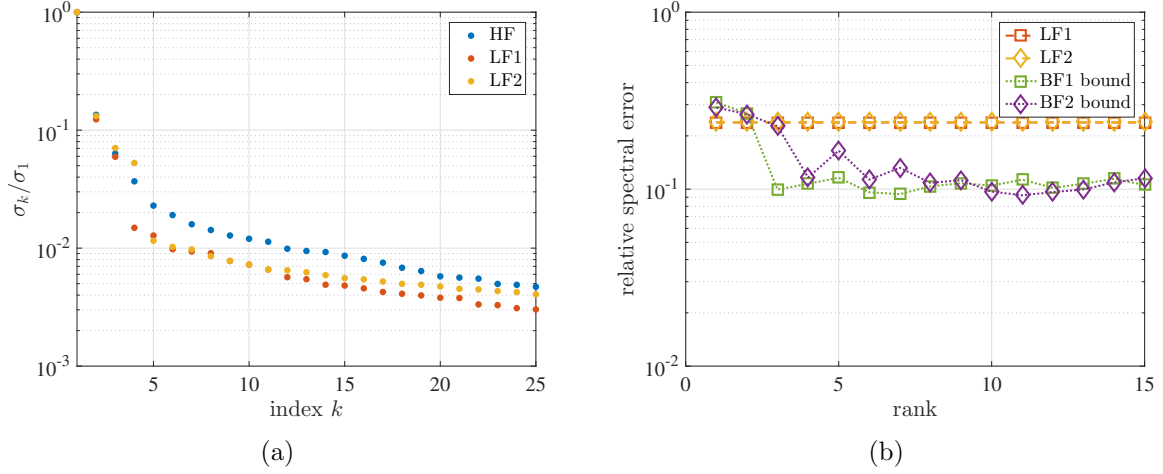


Figure 3.5: (a) Decay of normalized singular values of LF and HF matrices, using available data. (b) Error bound estimates for both BF approximations as a function of rank  $r$ . For comparison, relative spectral error of the LF data are provided.

3.5 (b) displays the error bound estimate of the BF approximation matrix as a function of rank  $r$ . Note, the number of simulations to calculate  $\hat{\epsilon}(\tau)$  is set to  $R = r + 2$ ; a more thorough assessment of the error bound will be discussed shortly. The error bound estimate levels out for values  $r \geq 4$ . For comparison, the LF1 and LF2 spectral errors are provided (see (3.22)), indicating that very few HF samples are needed to observe an improvement over the LF models. From these two results, a BF rank of  $r = 6$  is selected. The presentation of corresponding cost analysis is postponed until Section 3.5.3.

While the previous results indicate that the rank  $r = 6$  BF approximation more accurately describes the HF data, the calculated error must be verified by more thoroughly investigating the theoretical error bound estimates from Section 3.4.2. Specifically, the bound must be estimated for multiple values of  $R$ , the number of samples used to estimate  $\hat{\epsilon}(\tau)$ . Figure 3.6 provides the error bound estimates as a function of  $R$ , for (a) the BF1 model and (b) the BF2 model. Each point represents an error bound estimate calculated from  $R$  random columns of  $\mathbf{U}_H$  and  $\mathbf{U}_L$ , out of 26 total columns (thus not completely independent). The solid line represents the average value of the points at each value of  $R$ . Numerical results of [62] suggest a value of  $R \approx 2r$  will provide a

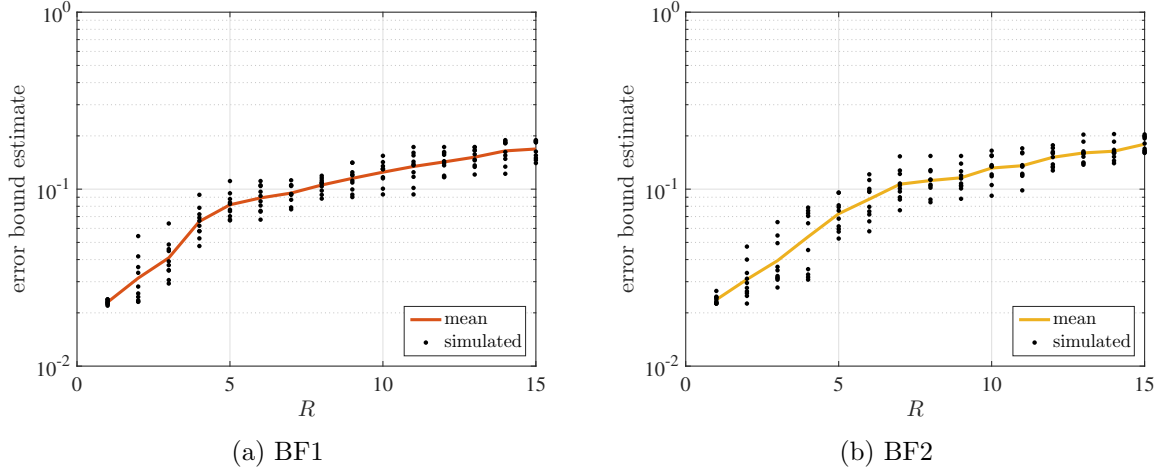


Figure 3.6: Error bound estimation in (3.19) using  $\hat{e}(\tau)$  and rank  $r = 6$  for varying values of  $R$ , for (a) BF1 and (b) BF2 models. Values are based on 10 different sets of  $R$  columns, where column selection not fully independent due to restricted number of HF samples.

true error bound. With rank  $r = 6$ , these results estimate the error bound to be 0.12 for both BF models. Recall from Figure 3.5 (b) that this error bound estimate is smaller than that of either LF model, indicating that improvement in accuracy may be estimated without knowledge of the true BF error. To compare the performances of the LF and BF approximations with regards to the heat flux QoI, consider the results of Figure 3.7. Figure 3.7 provides 17 simulated values of the heat flux QoI when using the HF, LF and BF models, and comparing only non-basis data. For each simulated value, the BF approximations are significantly more accurate than their LF counterparts. With regards to the relative  $\ell_2$  error, the BF QoIs are about  $7\times$  more accurate than the LF QoIs, with both BF1 and BF2 having errors of 0.03, LF1 an error of 0.20, and LF2 an error of 0.22.

To estimate the moments and PDF, as well as perform sensitivity analysis, a sparse PCE surrogate is formed from available data, as discussed in Section 3.4.3. Table 3.3 provides the QoI mean and coefficient of variation (CoV) values determined from the approximate coefficients of each LF and BF surrogate model, as well as the relative error between each QoI mean and the HF QoI mean. Note the HF QoI mean is calculated directly from available data. These results show that the BF1 mean is about  $100\times$  more accurate than the LF mean values and the BF2 mean is

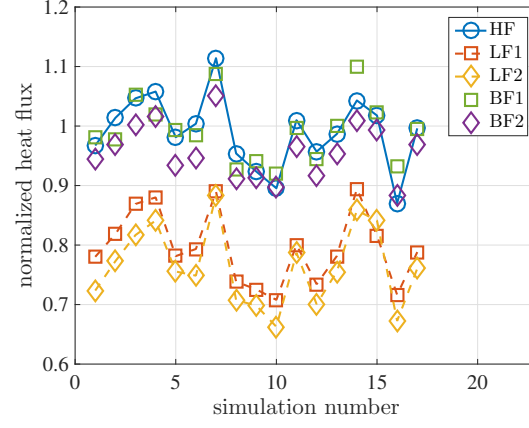


Figure 3.7: Normalized total heat flux values for 17 independent simulations, from the five different models, where the BF approximation is of rank  $r = 6$ . The BF approximations are more accurate with respect to the HF data than either LF QoI data. This data excludes simulations corresponding to basis data in the BF approximations.

Model	Mean	Rel.	CoV
Fidelity	QoI	Error	
LF1	0.80	20%	0.10
BF1	0.99	0.15%	0.07
LF2	0.78	22 %	0.11
BF2	0.97	2.7%	0.06
HF	1.0	-	0.08

Table 3.3: Statistics from Sparse PCE

about  $10\times$  more accurate than the LF mean values, where the BF1 model predicts within 0.15% of the HF QoI mean. The LF models, on the other hand, predict to within only 20 – 22% of the HF QoI mean. The CoV is provided for all models as well. The QoI variation for each fidelity model is comparable to that of the HF model, with that of BF1 most closely describing the HF data. Since the BF2 CoV is of the same order of magnitude as the respective error, it is not necessarily distinguishable and thus can't completely be relied on as an estimate for the HF CoV. The BF1 error, on the other hand, is significantly smaller than the CoV, and thus the CoV estimate can be trusted, indicating that the BF1 approximation most closely represents the HF data. Since the error and CoV values are small for both BF approximations, there is confidence in the error of mean estimation.

To estimate the QoI PDFs, histograms of the LF and BF data are generated from 100,000 sparse PCE samples. Figure 3.8 provides these normalized histograms for the heat flux QoI to compare with available HF data. Figure 3.8 (a) provides data derived from the LF1 and BF1 models, while Figure 3.8 (b) provides the histograms from LF2 and BF2 data. From previous results regarding the moment estimations, it was found that the CoV estimate of BF1 was reliable while that of the BF2 was not. An accurate CoV, combined with a small error indicates that the BF1 histogram is the most accurate PDF estimate of the HF data. In contrast, the LF surrogates had a large error, and thus the corresponding histograms are not dependable estimates of the HF PDF.

For the final set of results corresponding to the PCE coefficients, the Sobol' indices are calculated from the sparse PCE coefficients derived from the LF and BF data. The relative contribution of each parameter to the total sparse PCE variation is displayed in Figure 3.9 (a)-(d). The Sobol' indices determined from all four sets of sparse PCE coefficients clearly indicate that the uncertain parameters of heat fluxes from the radiated wall ( $\xi_{12}$ ) and opposite wall ( $\xi_{13}$ ) to the fluid are the most important with respect to the heat flux QoI. Comparatively, the remaining uncertain parameters are best distinguished from the BF1 results of Figure 3.9 (b) based on its accurate CoV estimate. This decomposition indicates that inputs  $\xi_i$  with  $i = 4 - 9, 11$  are next in terms of

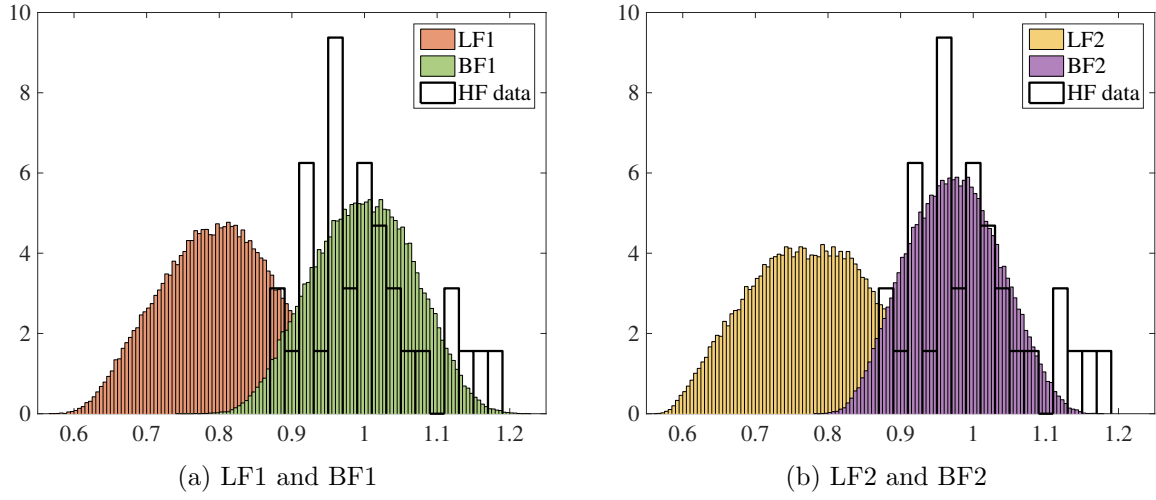


Figure 3.8: Normalized histogram of the total normalized heat flux through the  $x = \Delta 0.3L$  plane based on sparse PCE of (a) LF1 and BF1 QoI data, and (b) LF2 and BF2 QoI data.

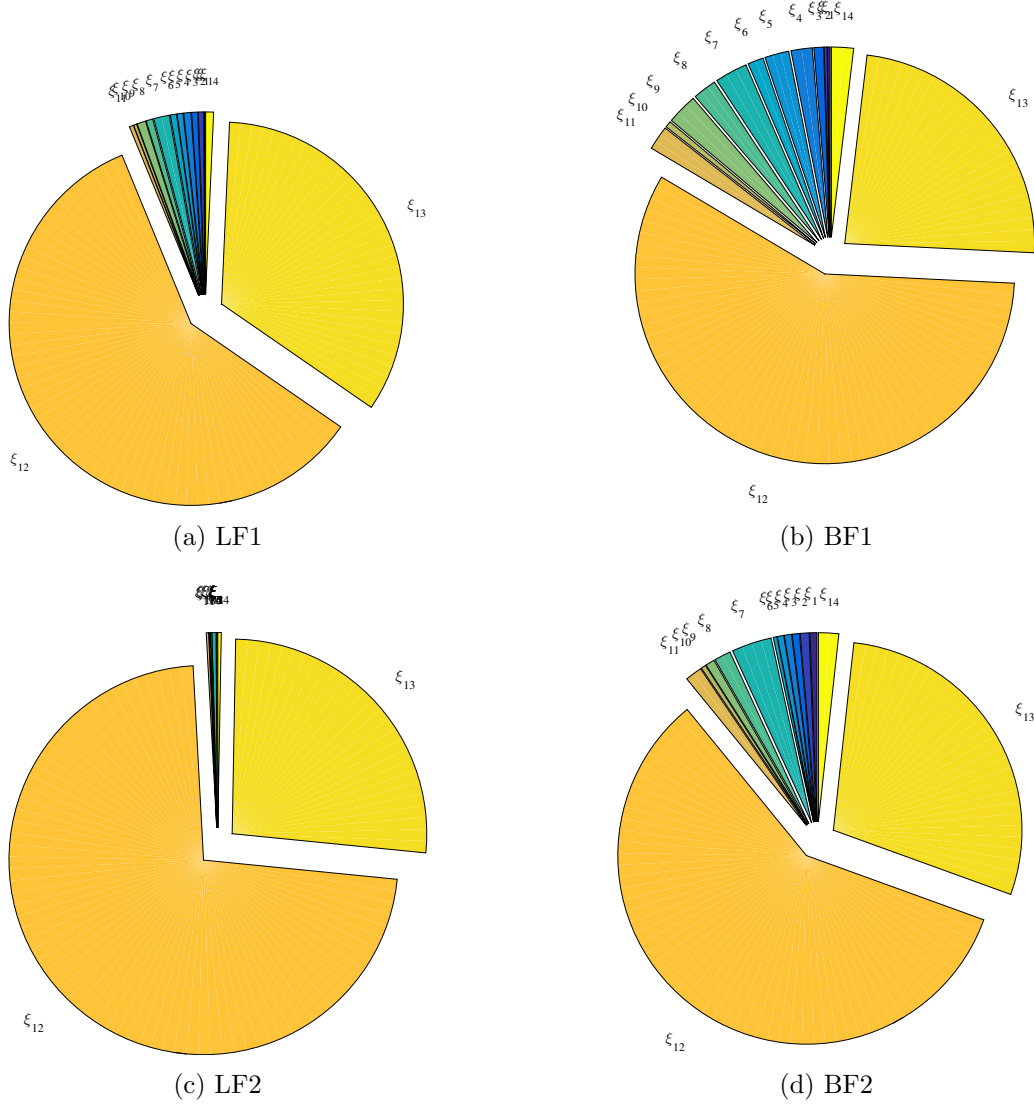


Figure 3.9: Importance of input parameters calculated from sparse PCE coefficients on the (a) LF1, (b) BF1, (c) LF2, and (d) BF2 heat flux data. Starting with parameter  $\xi_1$  at the top position, importance of each  $\xi_i$  is provided in counterclockwise order with respect to increasing  $i$ , with corresponding description provided in Table 3.2. Heat flux from the radiated wall ( $\xi_{12}$ ) and opposite wall ( $\xi_{13}$ ) to the fluid are the most important parameters as determined from all four model surrogates.

contribution to the total QoI variance, with about 1% of the variance, and  $\xi_i$  with  $i = 1 - 3, 10$  have the least important contribution, with  $< 0.1\%$  of the total variance. As the CoV estimates of the BF2 and LF surrogates were imprecise, further conclusions cannot be made from the corresponding results.

As the results of this section have shown, both BF approximations are significantly more accurate than either LF approximation, with the sparse PCE surrogate of the BF1 approximation most accurately describing the HF QoI data. To show that the BF approximation is a well justified reduced model, it must be shown, in addition, that there is significant computational savings compared to the HF model. A discussion of cost comparisons between the five models may be found in Section 3.5.3.

### 3.5.2 QoI # 2: Spatially-Averaged and Point Estimates of $\Delta T/T_0$ Along Profile at Probe Location

Next, time-averaged  $\Delta T/T_0 = (T - T_0)/T_0$  along the profile at the probe location is considered. The BF approximation employs realizations  $\mathbf{u}$  of the change in temperature  $\Delta T/T_0$  along this profile. After investigating the estimation of the full temperature profile for all models, focus is placed on estimating the spatial mean of  $\Delta T/T_0$  along the profile, as well as  $\Delta T/T_0$  at three points along the profile, namely,  $y/W = 0.5$ ,  $y/W = 0.1$ , and  $y/W = 0.05$  at the probe location.

Similar to the heat flux QoI, an optimal rank of the BF approximation must first be identified. As previously discussed, in Section 3.5.1, the ideal selected rank is the minimum value of  $r$  that captures the important modes of the solution. Doing so enables for improved accuracy, while restricting the number of HF simulations needed for the BF approximation, and thus limiting the associated computational cost of performing HF simulations. To aid in rank selection, consider the results of Figure 3.10. Figure 3.10 (a) provides the decay of singular values of LF and HF matrices, indicating that a rank  $r = 6$  approximation captures the larger modes of the LF data. Figure 3.10 (b) displays the BF approximation error bound estimation and calculated relative spectral errors (see (3.22)) for the LF data, as a function of rank  $r$ . For these error bound estimates,  $R = r + 2$

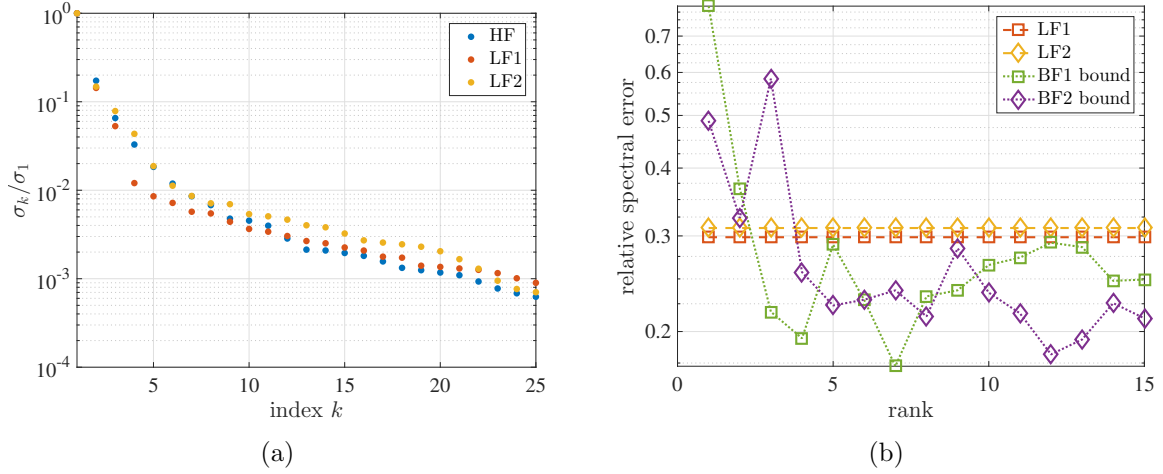


Figure 3.10: (a) Decay of normalized singular values of LF and HF matrices, using available  $\Delta T/T_0$  data along the profile at the probe location. (b) Error bound estimates of the rank  $r$  BF approximations for  $\Delta T/T_0$  QoI, as a function of  $r$ . For comparison, the LF1 and LF2 relative spectral errors are included as well.

simulations are used to estimate the values of  $\hat{\epsilon}(\tau)$ . As the error bound estimate levels out for values of  $r \geq 4$ , it is determined that the accuracy of the approximation requires at least rank  $r = 4$ . In comparison with the LF results, the error bound estimate shows that the error will be smaller than that of the LF approximations for these larger values of  $r$ . For the remaining results the BF approximation rank is fixed to  $r = 6$ .

Next a thorough investigation of the error bound is completed, with fixed rank  $r = 6$ . This is necessary as the theoretical error bound requires simulated values to estimate the bound for a range of  $R$  values. Figure 3.11 provides the error bound estimates for (a) the BF1 model and (b) the BF2 model, as a function of  $R$ . Single points indicate individual error bound calculations from  $R$  random columns of  $\mathbf{U}_H$  and  $\mathbf{U}_L$ , out of 26 total columns (and thus not completely independent). The solid line provides an average of these 10 points at each value of  $R$ . With rank  $r = 6$ , these results estimate the error bound to be 0.30 for both the BF1 approximation and the BF2 approximation. Compared to the error estimates of the LF models in Figure 3.10 (b), these error bound estimates indicate that the BF approximations will have an error less than or equal to that of the LF models; however, significant error improvement is not guaranteed, as in Section 3.5.1 for the heat flux data.



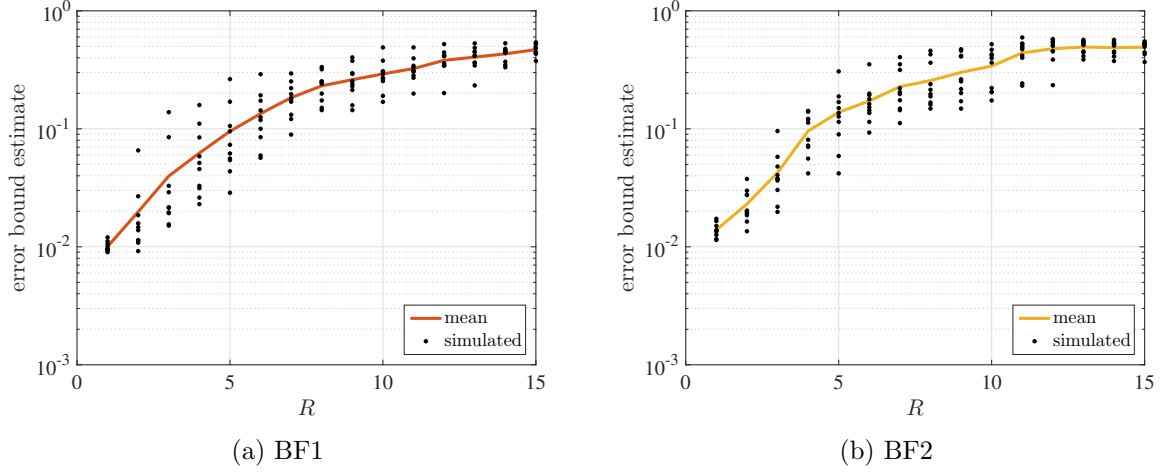


Figure 3.11: Error bound estimation in (3.19) using  $\hat{e}(\tau)$  and rank  $r = 6$  for varying values of  $R$ , for (a) BF1 and (b) BF2 models. Values are based on 10 different sets of  $R$  columns, where column selection not fully independent due to restricted number of HF samples.

To compare the abilities of the LF and BF models to reconstruct the  $\Delta T/T_0$  temperature profile at the probe location, the results from Figure 3.12 are considered. Figure 3.12 (a) displays the average  $\Delta T/T_0$  temperature profile derived from the LF, BF, and HF models. At most points along the profile, the mean BF1 and BF2  $\Delta T/T_0$  are observed to be more accurate than the mean of either LF model. To quantify this error, Figure 3.12 (b) provides the  $\ell_2$  error evaluated at each point along the profile (see (3.23)). While the BF models are more accurate than the LF models at most points, the interior of the profile exhibits the greatest improvement. Of most interest for UQ analysis are the  $\Delta T/T_0$  values near the walls and middle of the profile. As such, the focus is placed on four different QoIs of  $\Delta T/T_0$ : the spatial mean of  $\Delta T$  along the profile, and  $\Delta T/T_0$  quantities at points  $y/W = 0.5$ ,  $y/W = 0.1$ , and  $y/W = 0.05$  along the profile.

To gauge the performance at each of the four  $\Delta T/T_0$  QoIs, the results from Figure 3.13 are considered. Figure 3.13 (a)-(d) provides the simulated values of the four  $\Delta T/T_0$  QoIs when using the HF, LF, and BF models. Note these do not include the data used for the BF approximation basis. Figure 3.13 (a) and (b), which provide the mean value of  $\Delta T/T_0$  and the value of  $\Delta T/T_0$  at

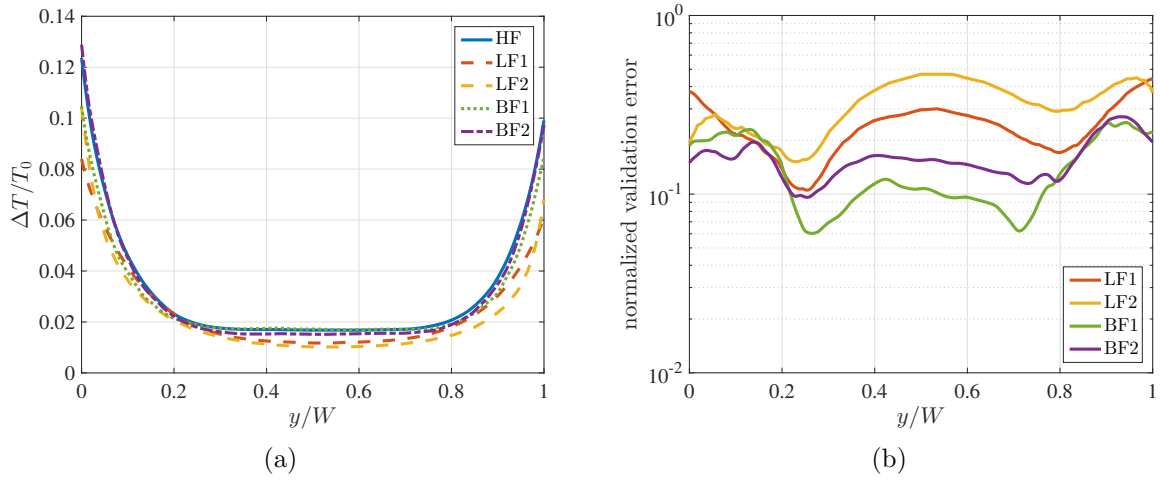
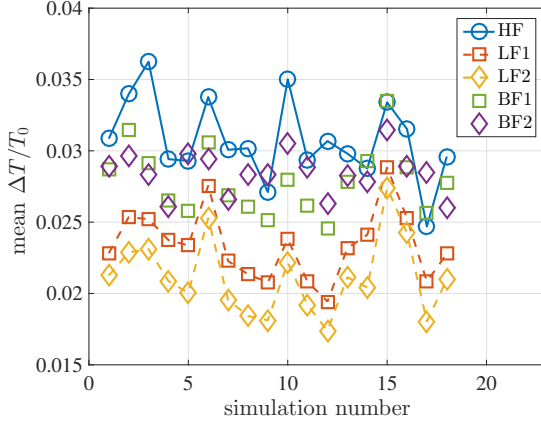
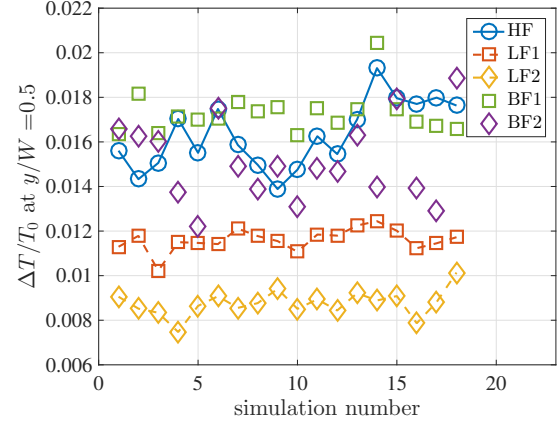


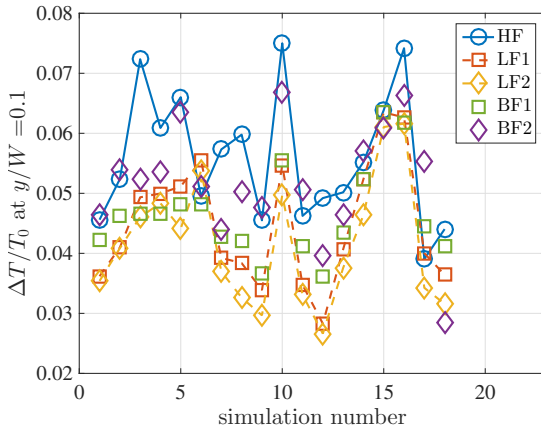
Figure 3.12: (a) Average  $\Delta T/T_0$  profile calculated from available simulations for all five models, with rank  $r = 6$  for the BF approximations. (b)  $\ell_2$  error estimates of  $\Delta T/T_0$  at each point along the profile.



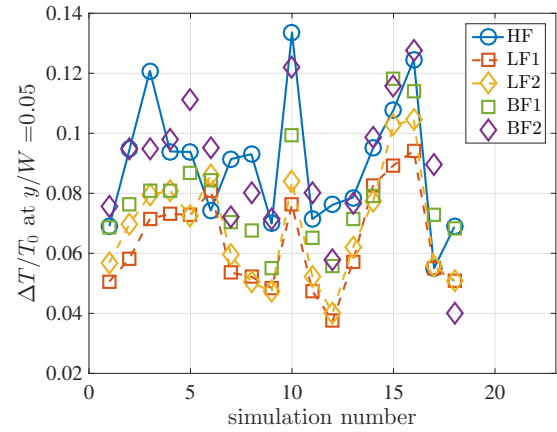
(a)



(b)



(c)



(d)

Figure 3.13:  $\Delta T/T_0$  values of 17 independent simulations for the QoI of (a) spatial mean along profile at probe location, and points (b)  $y/W = 0.5$ , (c)  $y/W = 0.1$ , and (d)  $y/W = 0.05$ , along probe profile. Simulated values are from the five different models, where the BF approximation is of rank  $r = 6$ .

$y/W = 0.5$ , respectively, show improved performance of the BF approximation compared to the LF models for all simulations. With regards to the  $\ell_2$  error of the QoIs, the mean values are about  $2\times$  more accurate and the values at  $y/W = 0.5$  are  $3\times$  more accurate. On the other hand, in Figure 3.13 (c) and (d), which provide the values of  $\Delta T/T_0$  at  $y/W = 0.1$  and  $y/W = 0.05$ , respectively, error improvement does not appear to be significant. For both QoIs, the BF approximations have a smaller  $\ell_2$  error than their corresponding LF models, however, the gain is smaller than  $2\times$ . These results suggest that the approximation performs well for the mean and  $\Delta T/T_0$  QoIs near the interior of the profile, but accuracy decays for values closer to the walls.

While prior temperature results were calculated from simulated values, to estimate the moments and PDFs, as well as perform sensitivity analysis, surrogates of the LF and BF models are formed via sparse PCE approximations. For each of the four temperature QoIs, the mean and CoV estimates determined from the sparse PCE coefficients are provided in Table 3.4 (a)-(d), as well as the relative error between each QoI mean and the HF QoI mean. For each subtable in Table 3.4, the BF mean estimates are more accurate than either of the LF estimates, by a factor of  $3 - 11\times$ . Between the two BF approximations, BF2 consistently has a small error. Aside from Table 3.4 (b), in which BF1 is more accurate (2% error vs. 4% error), the BF2 errors are the smallest for each subtable. In terms of the QoI CoV estimates, the values of all models are comparable to that of the HF data. In Table 3.4 (a) and (b) the CoV results are similar to those of the heat flux results; specifically, since the BF CoV is the same order of magnitude of the respective errors, it is not necessarily distinguishable and thus can't completely be relied on as an estimate for the HF CoV. However, as the BF approximations exhibit small errors and CoV values, the mean estimates are reliable. On the other hand, the BF CoV estimates of Table 3.4 (c) and (d) are larger than the corresponding error, and therefore are more dependable than the estimates in Table 3.4 (a) and (b). Because of this, stronger conclusions may be made when performing global sensitivity analysis.

To estimate the PDFs of the four  $\Delta T/T_0$  QoIs, histograms of the LF and BF sparse PCE surrogates are formed via 25,000 samples. These results are provided in Figure 3.14, where the left column results are associated with LF1 and BF1 surrogates, and the right column results are

Model	Mean	Rel.	CoV
Fidelity	QoI	Error	
LF1	0.024	21.7%	0.12
BF1	0.028	8.11%	0.11
LF2	0.022	28 %	0.13
BF2	0.029	3.3%	0.07
HF	0.031	-	0.10

(a) Spatial mean  $\Delta T/T_0$ 

Model	Mean	Rel.	CoV
Fidelity	QoI	Error	
LF1	0.046	16.1%	0.25
BF1	0.047	14.7%	0.21
LF2	0.042	24.0 %	0.27
BF2	0.052	5.1%	0.20
HF	0.055	-	0.20

(c)  $\Delta T/T_0$  at  $y/W = 0.1$ 

Model	Mean	Rel.	CoV
Fidelity	QoI	Error	
LF1	0.012	31%	0.04
BF1	0.017	2.0%	0.06
LF2	0.010	40 %	0.09
BF2	0.016	4.4%	0.10
HF	0.017	-	0.08

(b)  $\Delta T/T_0$  at  $y/W = 0.5$ 

Model	Mean	Rel.	CoV
Fidelity	QoI	Error	
LF1	0.066	27.3%	0.28
BF1	0.081	11.1%	0.24
LF2	0.071	22.6 %	0.30
BF2	0.088	2.8%	0.27
HF	0.091	-	0.25

(d)  $\Delta T/T_0$  at  $y/W = 0.05$ 

Table 3.4: Statistics from sparse PCE

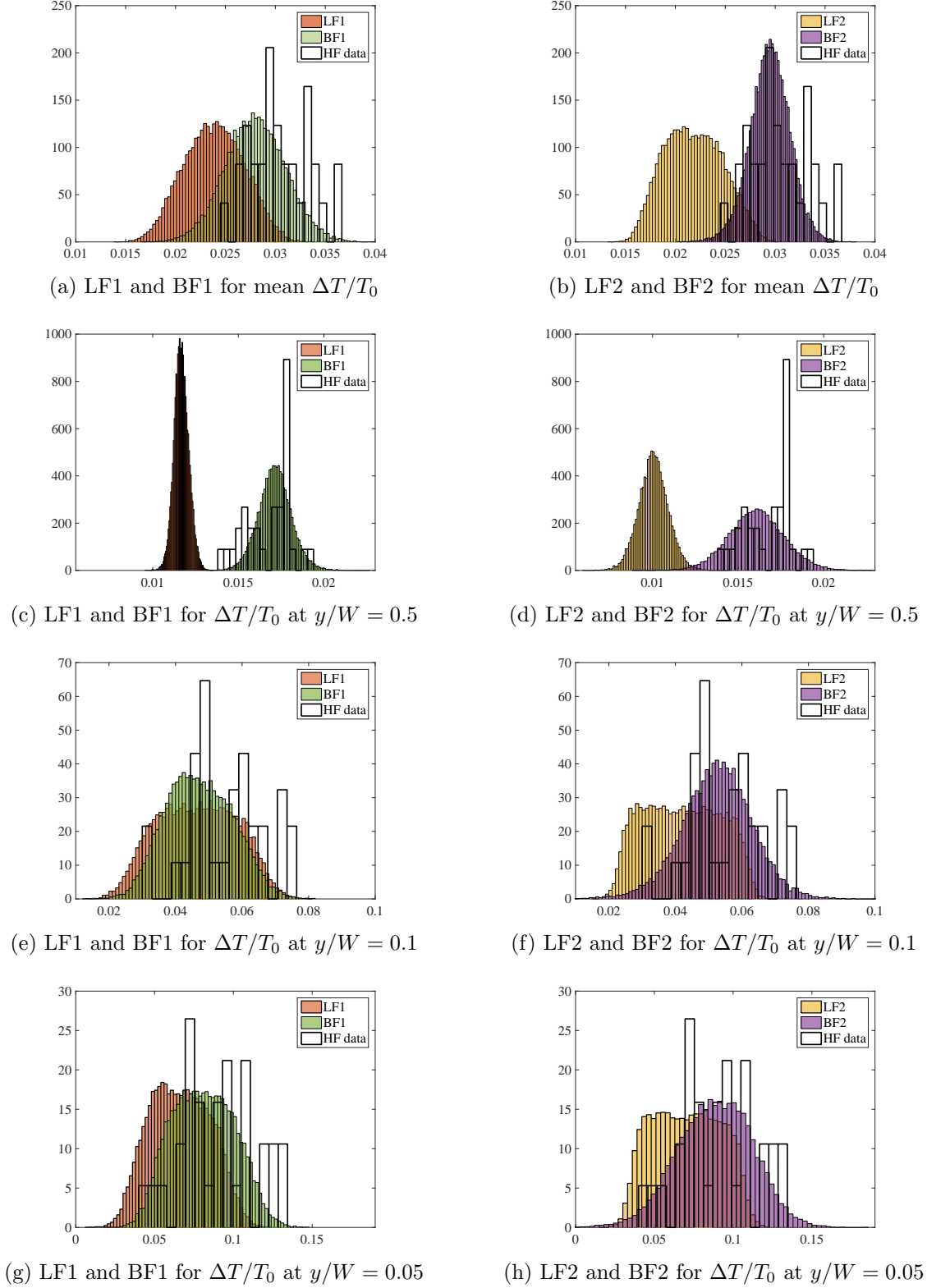


Figure 3.14: Normalized histograms of the LF1 and BF1 surrogate models (left column) and LF2 and BF2 surrogate models (right column) for the four  $\Delta T/T_0$  QoIs: mean ((a) and (b)),  $y/W = 0.5$  ((c) and (d)),  $y/W = 0.1$  ((e) and (f)), and  $y/W = 0.05$  ((g) and (h)). Histograms formed from 25,000 samples of the sparse PCE surrogates.

associated with LF2 and BF2 surrogates. For comparison available simulated HF data are provided as well. With the exception of Figure 3.14 (e) and (g), where the LF1 and BF1 results are closely overlaid, all of the histograms show the BF results more accurately approximating the HF data than the LF results. This significance is observed more so for Figure 3.14 (a)-(d), which displays the QoI histograms of the spatial mean of  $\Delta T/T_0$  and  $\Delta T/T_0$  at the center of the profile. In Figure 3.14 (c)-(h), which shows the  $\Delta T/T_0$  QoIs at specific points along the profile, the BF surrogates are consistently accurate. The LF surrogates are the least accurate in the center of the profile (Figure 3.14 (c) and (d)), and improve for QoI point estimates near the wall (Figure 3.14 (e)-(h)).

The last result considered is global sensitivity analysis via Sobol' indices, as calculated from the sparse PCE coefficients. From these estimates, two sets of results are presented: comparisons between the four models, and comparisons between the four QoIs. To compare the four models, sensitivity analysis is completed for the spatial mean  $\Delta T/T_0$  QoI. Figure 3.15 provides the decomposition of important parameters from the (a) LF1, (b) BF1, (c) LF2, and (d) BF2 models. All pie charts suggest that heat flux from the radiated wall to the fluid ( $\xi_{12}$ ) is the most important parameter affecting the QoI variance. However, it is important to note the results of Table 3.4 (a); specifically, the LF1 and LF2 errors are larger than the corresponding CoV estimates, indicating that the associated sensitivity analysis is not necessarily reliable. The BF1 and BF2 data, on the other hand, have an error that is smaller than the CoV, but on the same order of magnitude. As such, Figure 3.15 (b) and (c) show that the heat flux from the radiated wall and opposite wall to the fluid ( $\xi_{12}$  and  $\xi_{13}$ , respectively) are the two most important input parameters. Further conclusions cannot be made with regard to the remaining parameters as the CoV estimates of Table 3.4 (a) are of the same order as the error.

Based on the moment estimations of Table 3.4 and histograms of Figure 3.14, it is clear that the BF approximations provide better estimates of the HF QoIs than the corresponding LF models. As a consequence, Figure 3.16 provides the importance of input parameters of the four QoIs as determined by the BF2 sparse PCE coefficients. The BF2 model is selected as it consistently has a low error (see Table 3.4), and does so with lower cost than the BF1 surrogate model. Figure

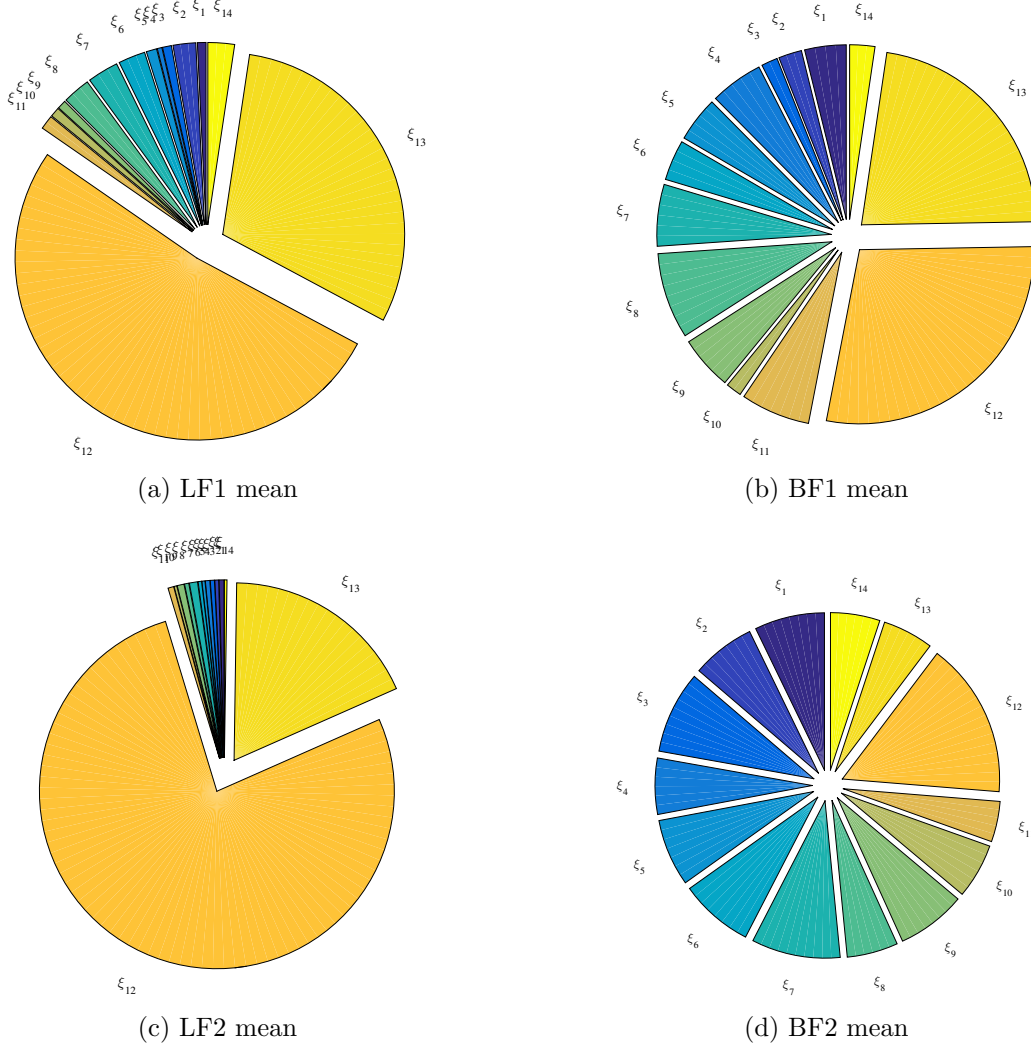


Figure 3.15: Importance of input parameters for the spatial mean  $\Delta T/T_0$  QoI from sparse PCE coefficients of the (a) LF1 (b) BF1, (c) LF2, and (d) BF2 surrogate models. Starting with parameter  $\xi_1$  at the top position, importance of each  $\xi_i$  is provided in counterclockwise order with respect to increasing  $i$ , with corresponding description provided in Table 3.2.



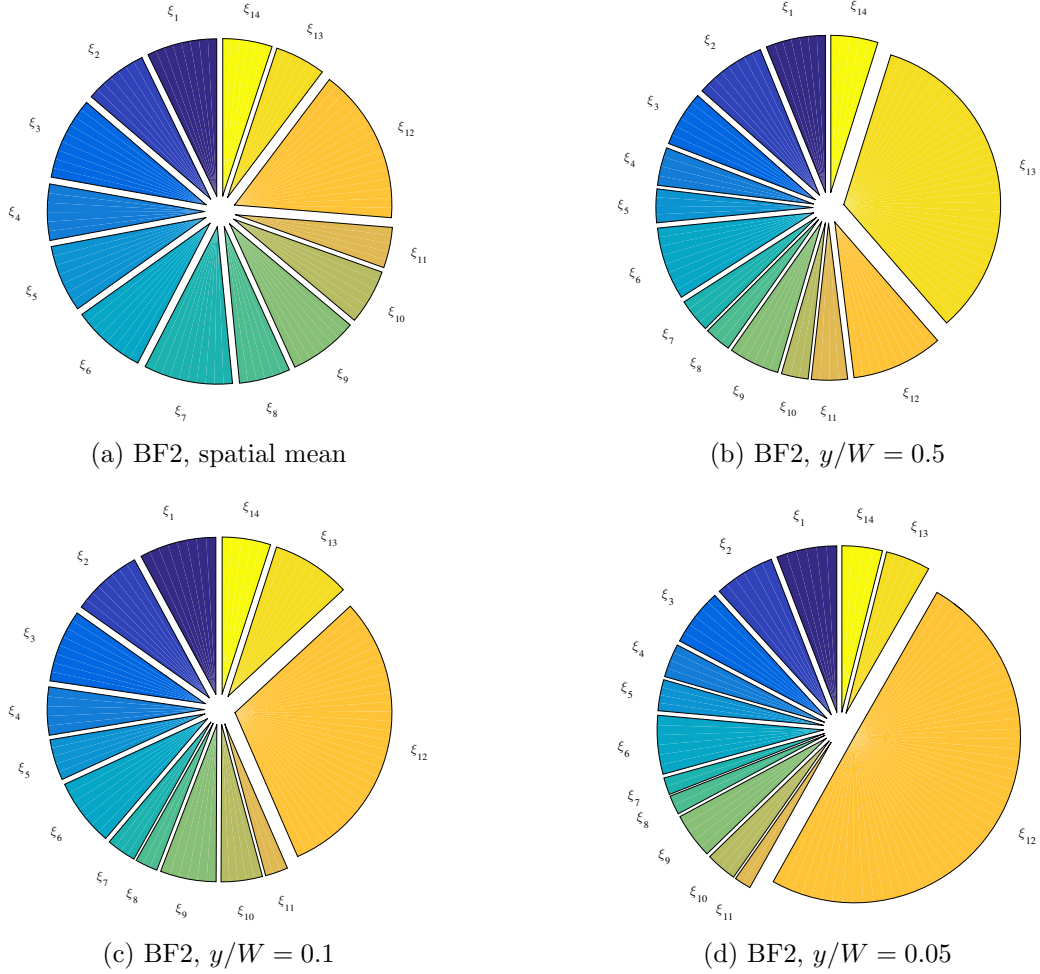


Figure 3.16: Importance of input parameters from sparse PCE coefficients from the BF2 model for the  $\Delta T/T_0$  QoIs of (a) spatial mean along profile at probe location (b)  $y/W = 0.5$  along profile at probe location, (c)  $y/W = 0.1$  along profile at probe location, and (d)  $y/W = 0.05$  along profile at probe location. Starting with parameter  $\xi_1$  at the top position, importance of each  $\xi_i$  is provided in counterclockwise order with respect to increasing  $i$ , with corresponding description provided in Table 3.2.

3.16 (a)-(d) shows how the importance of parameters changes with the four QoIs. From Figure 3.16 (a)-(d), the heat flux from the radiated wall to the fluid ( $\xi_{12}$ ) is the most important parameter for the mean  $\Delta T/T_0$  QoI, as well as the two point QoIs near the radiated wall ((a), (b), and (c), respectively); in contrast, heat flux from the opposite wall to the fluid ( $\xi_{13}$ ) is the most important parameter for the  $\Delta T/T_0$  QoI in the middle of the profile (see (d)). This suggests that, over the whole profile, variations in the heat flux from the radiated wall will greatly affect the  $\Delta T/T_0$  values, but more so at points close to the radiated wall. For points further from this wall, the variations in this heat flux will play less of a role in the variations of  $\Delta T/T_0$ .

In terms of the remaining parameters, the results of Table 3.4 (a) and (b) indicate that there is lack of sufficient precision in the variance estimate to conclude the importance of the remaining parameters for Figure 3.16 (a) and (b). However, the CoV estimates in 3.4 (c) and (d) are significantly larger than the corresponding error estimates for the BF2 models, allowing for further conclusions to be made from the sensitivity analysis results. Specifically, the data of Figure 3.16 (c) and (d) show that the remaining parameters contribute equally to the variance of the  $\Delta T/T_0$  QoIs near the radiated wall.

The  $\Delta T/T_0$  results of this section show that the BF approximations provide a more accurate reduced model than either of the LF models, with greatest improvements for mean  $\Delta T/T_0$  and  $\Delta T/T_0$  at the interior of the profile, in comparison to HF data. In addition, theoretical error results guarantee that the BF approximations will be at least as accurate as the corresponding LF data. As will be shown in the following section, the cost of this approximation is close to that of the LF models when many simulations are required.

### 3.5.3 Computational Cost Comparisons of the Five Models

The final component necessary to justify this BF approximation is to perform a cost comparison of all five models. Figure 3.17 provides the approximate number of core-hours needed to generate  $N$  converged simulations from the HF, LF, and rank  $r = 6$  BF models, extrapolated to large values of  $N$ . The number of simulations generated in this study are indicated by markers.

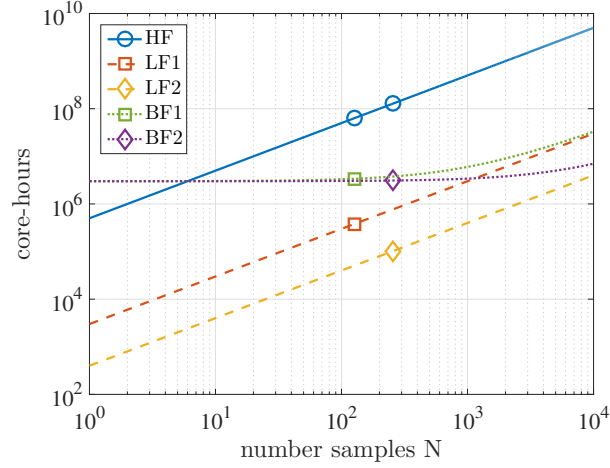


Figure 3.17: Number of core-hours on Mira (ALCF) [2] to obtain  $N$  simulated values from each model. Markers provide number of core-hours for  $N = 128$  and  $N = 256$ , to provide the cost of the simulated values generated for this study.

For the BF1, the cost to generate  $N = 128$  samples is  $20\times$  less expensive than the HF model, and for the BF2, the cost to generate  $N = 256$  samples is  $50\times$  less expensive than the HF model. In comparison, the LF1 and LF2 models are  $170\times$  and  $1300\times$  less expensive, respectively; however, as shown in the results, they are poor approximations to the HF data. As the  $r$  HF simulations greatly affect the cost of the BF approximation, significant cost reduction is observed for larger values of  $N$ . For  $\mathcal{O}(10^3)$  samples, which corresponds to values of  $N$  that are of interest in the context of the application studied in this work, the computational cost of the BF models more closely aligns with the cost of the LF models. While  $N = 10^3$  HF simulations is approximately 500M core-hours, obtaining the equivalent number of simulations is approximately 6M and 3.5M core-hours for the BF1 and BF2 approximations, respectively. This dramatic cost improvement, without as significant a loss of accuracy that is observed with the LF models, makes the BF approximation a powerful tool to perform UQ for large-scale problems.

### 3.6 Conclusions

As the results of this work show, when simulated solutions exhibit a low-rank structure, i.e., rapid decay of singular values, low-rank estimates such as this BF approximation may prove

successful in terms of accuracy and cost reduction. This is of great significance for complex, multi-physics applications, as performing UQ analysis requires many large-scale simulations, which is computationally infeasible when restricted only to the HF model. For the system studied in this work, focus was placed on estimating the statistics of two thermal QoIs related to time-averaged heat flux and time-averaged temperature near the outflow boundary. From early results in each section, it was shown that the estimated error bounds between the BF approximations and HF data were smaller than the error between associated LF and HF data. As an improvement in accuracy was established, generating approximate sparse PCE surrogates from fixed sample sizes enabled to further calculate QoI statistics via PDF estimation and performing sensitivity analysis to understand which parameters most affect variability in the QoIs.

Ongoing and future work focuses on exploring the performance of the BF approximation under more challenging physical conditions in terms of turbulent intensities, particle loading, and radiation intensity. Of notable interest is also the assessment of the approach in predicting higher order statistics, e.g., particles and temperature fluctuations, and more complex physical phenomena, such as particle clustering and turbulence modulation.

## Chapter 4

### A Low-rank Control Variate for Multilevel Monte Carlo Simulation of High-dimensional Uncertain Systems

#### 4.1 Introduction

The use of uncertainty quantification as a tool to assess the prediction accuracy of simulation models of physical systems has been increasing at a rapid rate over the last decade. By accounting for the uncertainties of input data in models, such as initial conditions, boundary conditions, or other physical parameters, the objective is to establish the predictive capabilities of simulations by quantifying the uncertainty in the quantities of interest (QoI's). To this end and within the probabilistic framework, several methods, e.g., polynomial chaos expansions [48, 136, 29] and stochastic collocation [137, 88], have been developed and proven successful in various applications. However, it is known that the computational cost of these methods grows rapidly as a function of the number of random variables describing model uncertainties, a phenomenon referred to as *curse of dimensionality*.

An alternative class of techniques rely on the Monte Carlo (MC) simulation or its variants, where the statistics of the QoI are estimated using an ensemble of (random) realizations of the QoI. The cost of such estimations, while may be prohibitive, is *formally* independent of the number of input variables. In details, let  $\boldsymbol{\xi} = (\xi_1, \dots, \xi_d)$  denote the  $d$ -vector of random variables, with joint probability density function  $\rho_{\boldsymbol{\xi}}(\boldsymbol{\xi})$ , representing the uncertainty in the inputs. Let  $Q = Q(\boldsymbol{\xi})$  denote a scalar-valued QoI depending on  $\boldsymbol{\xi}$  and  $Q_M$  its approximation obtained via simulation. The subscript  $M$  denotes the number of deterministic degrees of freedom, e.g., number of grid points in

a finite element model, controlling the accuracy of  $Q_M$  relative to  $Q$ . The goal is to approximate the statistics of  $Q$ , e.g., the mean of  $Q$ ,  $\mathbb{E}[Q]$ , using the realizations of  $Q_M$ . Given a set of  $N$  samples of inputs, each denoted by  $\boldsymbol{\xi}^{(i)}$  and drawn according to  $\rho_{\boldsymbol{\xi}}(\boldsymbol{\xi})$ , and the corresponding realizations of  $Q_M$ , given by  $Q_M^{(i)} = Q_M(\boldsymbol{\xi}^{(i)})$ , the MC approximation of  $\mathbb{E}[Q]$  is

$$\mathbb{E}[Q] \approx \mathbb{E}[Q_M] \approx \hat{Q}_{M,N}^{MC} = \frac{1}{N} \sum_{i=1}^N Q_M^{(i)}. \quad (4.1)$$

Following the notation in [24],  $\hat{Q}_{M,N}^{MC}$  in (4.1) is the MC estimator of  $\mathbb{E}[Q_M]$  using  $N$  samples of  $Q_M$  with the Mean Square Error (MSE)

$$\text{MSE}(\hat{Q}_{M,N}^{MC}, \mathbb{E}[Q]) = \frac{1}{N} \mathbb{V}[Q_M] + (\mathbb{E}[Q_M - Q])^2, \quad (4.2)$$

where  $\mathbb{V}$  is the variance operator and  $\text{MSE}(\hat{Q}_{M,N}^{MC}, \mathbb{E}[Q])$  denotes the MSE of  $\hat{Q}_{M,N}^{MC}$  with respect to  $\mathbb{E}[Q]$ . We note that, in this paper, the hat operator indicates the MC estimator of the corresponding expectation. In (4.2), the MSE is decomposed into the *sampling error*  $\frac{1}{N} \mathbb{V}[Q_M]$ , controlled by the variance of  $Q_M$  and the number of samples, and the *discretization error*  $(\mathbb{E}[Q_M - Q])^2$ , which measures how closely the model simulates the true solution. As can be seen from (4.2), the sampling error decays slowly as a function of  $N$ , but with a rate that is independent of the dimension  $d$ , implying that the standard MC simulation does not formally suffer from the curse of dimensionality.

Aside from the necessary refinement in the model to reduce the discretization error, there are only two options to improve the MSE of a MC estimate: increasing the sample size  $N$  or using a variance reduction technique. Due to the cost incurred by the first option, it is more practical to consider the use of a variance reduction technique, such as importance sampling or control variates (CV) [8]. In particular, the CV approach considers a second quantity  $Z$ , such that it is correlated with  $Q_M$ , is cheaper than  $Q_M$  to evaluate, and whose expectation is either known or can be approximated with relatively small cost. Then a new variable,

$$W = Q_M - \theta(Z - \mathbb{E}[Z]),$$

is constructed that has the same mean as  $Q_M$ , i.e.,  $\mathbb{E}[W] = \mathbb{E}[Q_M]$ , thus suggesting the use of MC estimate of  $\mathbb{E}[W]$  as a proxy for  $\mathbb{E}[Q_M]$ . In doing so, the gain is that, depending on the choice of

$\theta$ , the MC estimator of  $\mathbb{E}[W]$  features a reduced MSE (or variance). Stated differently, a smaller number of  $W$  realizations, hence  $Q_M$  realizations, are needed for a comparable MSE when CV is applied. For scenarios when  $Z$  is poorly correlated with  $Q_M$ , a notable MSE reduction is not observed. If, in addition, the cost of estimating  $\mathbb{E}[Z]$  is large, it is likely that this CV will not result in a cost improvement over standard MC.

Multilevel Monte Carlo (MLMC), proposed in [65, 49], is a generalization of CV, which constructs a sequence of control variates  $Z$  based on approximations of  $Q$  on a set of models that are cheaper to simulate than the one for  $Q_M$ , hence the term *multilevel*. A common example of a cheaper model is to approximate  $Q$  on coarser grids with number of degrees of freedom smaller than  $M$ . While the notion of levels can go beyond a grid-based construction, we limit the scope of this study to such an approach. For the interest of a simpler introduction, we delay the full presentation of MLMC to Section 4.2, and instead focus on the two-level formulation next.

Taking  $\theta = 1$  and  $Z = Q_m$ , with  $m < M$ , to be the QoI approximated from a coarser grid than that of  $Q_M$ , the two-level MLMC variable is given by  $W = Q_M - (Q_m - \mathbb{E}[Q_m]) = \mathbb{E}[Q_m] + (Q_M - Q_m)$ , with expected value

$$\begin{aligned}\mathbb{E}[W] &= \mathbb{E}[Q_m] + \mathbb{E}[Q_M - Q_m] \\ &= \mathbb{E}[Q_M].\end{aligned}\tag{4.3}$$

To approximate  $\mathbb{E}[W]$ , or equivalently  $\mathbb{E}[Q_M]$ , MC is applied independently to the two expectations in the right-hand-side of the first equation in (4.3),

$$\hat{W} = \frac{1}{N_m} \sum_{i=1}^{N_m} Q_m^{(i)} + \frac{1}{N_M} \sum_{i=1}^{N_M} (Q_M^{(i)} - Q_m^{(i)}).\tag{4.4}$$

As compared to the standard MC estimator of  $\mathbb{E}[Q_M]$  given in (4.1), the estimation of  $\mathbb{E}[Q_m]$  in (4.4) also involves drawing samples of  $Q_m$ , which are less expensive. More importantly, when  $Q_m$  is close to  $Q_M$ , estimating  $\mathbb{E}[Q_M - Q_m]$  requires fewer samples of  $Q_M$ , as  $(Q_M - Q_m)$  features a smaller variance. In practice, depending on the cost of simulating the two models as well as the variances of  $Q_m$  and  $(Q_M - Q_m)$ , the numbers of samples of  $Q_m$ ,  $N_m$ , and  $Q_M$ ,  $N_M$ , are selected

such that the overall estimation cost, for a given accuracy, is minimal. MLMC expands upon this concept by including multiple levels, as delineated in Section 4.2.

Very recently the combination of MLMC and CV, referred to as *multilevel control variates* (MLCV), has been used with the aim of further reducing the variance of  $(Q_M - Q_m)$  and thereby improving the computational cost [93]. As an approach to MLCV, in [93], the authors consider the solution to a linear diffusion problem, where the diffusion coefficient is modeled by a rough random field represented by a large number of independent random variables. An auxiliary diffusion problem with smoothed (low-dimensional) coefficient is developed, whose solution – computed via stochastic collocation – is employed as the control variate for the MLMC simulation of the original problem. This particular construction of the control variate relies on the accuracy of the smoothed problem in predicting the solution to the original one. Other related work [124, 131] establishes multilevel control variates within a MC simulation framework. For instance, in [131], a model reduction approach based on the reduced basis method, see, e.g., [114], is considered for the case of linear, stochastic elliptic PDEs. In particular, in an off-line stage, a set of basis functions, i.e., reduced basis, is identified from the realizations of the solution on a target mesh. The PDE solution is then approximated in an increasing number of reduced basis functions, via a discontinuous Galerkin formulation, which are then employed as control variates. The levels here are defined based on the size of the reduced basis used, as opposed to the number of grid points in the spatial discretization of the PDE.

#### 4.1.1 Contributions of This Work

In this work we propose a different construction of a multilevel control variate that may be applicable to more general classes of problems. That is, our method may be applied to general nonlinear equations and is not limited to diffusion-type equations. In addition, we are able to use a standard PDE solver, rather than a different solver based on reduced basis, to generate the control variate and solution samples. The main idea behind this approach is to create a control variate  $Z$  for  $(Q_M - Q_m)$  that is obtained based on a low-rank approximation of  $Q_M$ , using a *small* set



of selected samples of fine grid solution identified from the realizations of the coarse grid solution. This low-rank approximation of the fine grid solution, based on ideas in [91, 142, 28] and following the algorithm described in [27], consists of three main steps. In the first step, a reduced basis, say of size  $r \ll M$ , is identified for the coarse grid solution, together with an interpolation rule that gives an arbitrary realization of the coarse grid solution in that basis. For this purpose, we borrow ideas from matrix interpolative decomposition (MID) as presented in [23]. The second step entails the identification of a reduced basis for the fine grid solution. For this, the fine counterpart to the coarse grid reduced basis is generated, which requires  $r$  realizations of the fine grid solution. In the third step, the same interpolation rule, as for the coarse grid solution, is applied to generate the low-rank approximation of its fine counterpart (and subsequently  $Q_M$ ). Based on this construction, we are able to formulate our low-rank control variate  $Z$  for  $(Q_M - Q_m)$ , and then the control variate equation

$$W = (Q_M - Q_m) - \theta(Z - \mathbb{E}[Z])$$

is used to estimate  $\mathbb{E}[Q_M - Q_m]$ , which requires setting  $\theta$  and estimating  $\mathbb{E}[Z]$ . For these, we consider work from [98], where  $\theta$  is chosen such that the MSE of  $\hat{W}$ , with respect to  $\mathbb{E}[Q_M - Q_m]$ , is minimized. In addition,  $\mathbb{E}[Z]$  is estimated primarily from an independent MC simulation of the coarse grid solution.

An important feature of this MLCV approach is worthwhile highlighting. Depending on the rank of the fine model solution and the cost of simulating  $Q_m$ , relative to  $Q_M$ ,  $Z$  may be sampled with a cost that is considerably smaller than that of  $Q_M$ . This, together with the MSE reduction achieved by  $\hat{W}$ , will lead to a smaller number of samples of  $Q_M$  required for a similar accuracy as in standard MLMC.

This paper is organized into four sections. A background on MLMC is provided in Section 4.2, including the setup and improvements it provides over MC, as well as a discussion of MLMC theory. Section 4.3 focuses on the formulation of the MLCV method of this paper, including the construction of the control variate and the MLCV estimator such that the MSE is minimized. This section also

includes an algorithm to outline the process, as well as a brief discussion regarding considerations to take into account when taking this MLCV approach. In Section 4.4 the numerical results of this new method as well as comparison with MLMC are presented using two test problems. The first test considers an eigenvalue problem associated with a linear elasticity problem in an L-shaped domain, followed by the second case, a thermally driven flow problem in a square domain.

## 4.2 Background on Multilevel Monte Carlo

From [65, 49], we next describe the use of MLMC as a variance reduction technique for the MC simulation of differential equations with random data. Consider a sequence of spatial discretizations, i.e., grids, for the governing equations with increasing accuracy, where each discretization is indexed by a *level* parameter  $\ell$ ,  $\ell = 0, \dots, L$ . Let  $M_\ell$  denote the number of degrees of freedom in the level  $\ell$  discretization model, such that  $M_0 < M_1 < \dots < M_L$ . Unlike in MC, MLMC does not directly approximate the expectation of QoI on the finest (target) level  $L$ . Instead, it uses the telescoping sum

$$\mathbb{E}[Q] \approx \mathbb{E}[Q_L] = \mathbb{E}[Q_0] + \sum_{\ell=1}^L \mathbb{E}[Q_\ell - Q_{\ell-1}], \quad (4.5)$$

which is the expected value of the QoI on the coarse level plus the sum of expectations of correction terms. Here,  $Q_\ell$  is the QoI determined from the level  $\ell$  grid with  $M_\ell$  degrees of freedom. Defining the multilevel correction variable

$$Y_\ell = Q_\ell - Q_{\ell-1}, \quad (4.6)$$

for  $\ell = 1, \dots, L$ , and  $Y_0 = Q_0$ , the sum in (4.5) can be rewritten as

$$\mathbb{E}[Q_L] = \sum_{\ell=0}^L \mathbb{E}[Y_\ell].$$

In MLMC, the expectation of each correction variable  $Y_\ell$  is independently computed with  $N_\ell$  MC samples of  $Y_\ell$ ,

$$\mathbb{E}[Y_\ell] \approx \hat{Y}_\ell = \frac{1}{N_\ell} \sum_{i=1}^{N_\ell} Y_\ell^{(i)} = \frac{1}{N_\ell} \sum_{i=1}^{N_\ell} (Q_\ell^{(i)} - Q_{\ell-1}^{(i)}), \quad (4.7)$$

where each  $Q_\ell^{(i)}$  and  $Q_{\ell-1}^{(i)}$  is generated by applying the same realization  $\boldsymbol{\xi}^{(i)}$  to the level  $\ell$  and  $\ell-1$  grids, respectively. Building on the notation once more, the MLMC estimator,  $\hat{Q}_L^{ML}$ , is defined to

be

$$\hat{Q}_L^{ML} = \sum_{\ell=0}^L \hat{Y}_\ell.$$

When  $Q_\ell$  converges to  $Q$  as a function of  $\ell$ , the variance of  $Y_\ell$  converges to zero. This allows for  $\hat{Y}_\ell$  to converge with fewer samples of  $Q_\ell$  for larger, and more expensive, levels  $\ell$ . To determine how to pick  $N_\ell$  such that the same level of accuracy as in MC is reached with a reduced computational cost, the MSE of  $\hat{Q}_L^{ML}$  must first be considered.

#### 4.2.1 MSE of MLMC

The MSE of the MLMC estimator, much like that of the MC estimator in (4.2), can be decomposed into the sampling and discretization errors, given by

$$\text{MSE}(\hat{Q}_L^{ML}, \mathbb{E}[Q]) = \sum_{\ell=0}^L \frac{1}{N_\ell} \mathbb{V}[Y_\ell] + (\mathbb{E}[Q_L - Q])^2, \quad (4.8)$$

where the sampling error of MLMC is a sum of MC sampling errors on each level. For a desired MSE bound of magnitude  $\varepsilon^2$ , the sampling error (as well as the discretization error) is ideally required to be bounded by  $\varepsilon^2/2$ ; that is,

$$\sum_{\ell=0}^L \frac{1}{N_\ell} \mathbb{V}[Y_\ell] \leq \frac{\varepsilon^2}{2}. \quad (4.9)$$

Provided an estimate of  $\mathbb{V}[Y_\ell]$ , (4.9) serves as a constraint to estimate the number of samples  $N_\ell$  drawn from each level.

#### 4.2.2 Number of Samples $N_\ell$ for $\hat{Y}_\ell$

In MLMC,  $N_\ell$  is determined such that the total cost, given by

$$\mathcal{C}(\hat{Q}_L^{ML}) = \sum_{\ell=0}^L N_\ell \mathcal{C}_\ell, \quad (4.10)$$

is minimized, where  $\mathcal{C}_\ell$  is the cost of generating a sample of  $Y_\ell$  and is given by

$$\mathcal{C}_\ell = \mathcal{C}(Y_\ell) = \mathcal{C}(Q_\ell) + \mathcal{C}(Q_{\ell-1}), \quad (4.11)$$

for  $\ell = 1, \dots, L$ , and  $\mathcal{C}_0 = \mathcal{C}(Q_0)$ . In practice, the average CPU time for obtaining a sample of  $Y_\ell$  can be used to determine  $\mathcal{C}_\ell$ . Alternatively, when the computational complexity of the adopted solvers with respect to  $M_\ell$  is known, a relation of the form

$$\mathcal{C}_\ell \lesssim M_\ell^\gamma, \quad \gamma > 0,$$

may be considered, where  $\lesssim$  denotes inequality up to a positive constant. By using the method of Lagrange multipliers to minimize the total cost in (4.10), with an  $\varepsilon^2/2$  sampling error constraint in (4.9), the ideal number of samples  $N_\ell$  from each level can be computed from

$$N_\ell \geq \frac{2}{\varepsilon^2} \left[ \sum_{k=0}^L \sqrt{\mathbb{V}[Y_k] \mathcal{C}_k} \right] \sqrt{\frac{\mathbb{V}[Y_\ell]}{\mathcal{C}_\ell}}, \quad (4.12)$$

see, e.g., [50]. In order to calculate the values of  $N_\ell$ , a *pilot run* of MLMC must first be implemented; that is, MLMC is completed on a relatively small sample size that may be constant across all levels. By doing so, the expectation, variance, and cost on each level can be estimated to determine the required number of levels and optimal  $N_\ell$  from (4.12), such that a given MSE tolerance can be achieved. Subsequently, a second run of MLMC is completed, this time using only the calculated sample size  $N_\ell$  for each level. To make the method as efficient as possible, simulations from the pilot run are incorporated into the second run of MLMC.

#### 4.2.3 MLMC Convergence Guarantee

In order to guarantee an  $\varepsilon^2$  bound on the MSE in (4.8), we turn to the MLMC theorem as presented in [24]. While the number of samples on each level can be determined such that the desired sampling error bound  $\varepsilon^2/2$  is attained, it does not imply the same for the discretization error. The MLMC theorem provides the constraints in order to maintain the prescribed MSE bound. Assuming that there exists an integer  $s > 1$  such that  $M_\ell = sM_{\ell-1}$ , for  $\ell = 1, \dots, L$ , and

$$(1) \quad |\mathbb{E}[Q_\ell - Q]| \lesssim M_\ell^{-\alpha},$$

$$(2) \quad \mathbb{V}[Y_\ell] \lesssim M_\ell^{-\beta},$$

$$(3) \mathcal{C}_\ell \lesssim M_\ell^\gamma,$$

for some constants  $\gamma, \beta > 0$  and  $\alpha \geq \frac{1}{2}\min(\beta, \gamma)$ , Theorem 1 in [24] states that there exists a positive integer  $L$  depending on  $\varepsilon$  and a sequence  $\{N_\ell\}_{\ell=0}^L$ , such that  $\text{MSE}(\hat{Q}_L^{ML}, \mathbb{E}[Q]) < \varepsilon^2$ . This relation indicates whether or not it is necessary to refine the model further to meet the desired discretization error. In Section 4.4, estimates of  $\alpha$  and  $\beta$  will be provided as an illustration that the theory holds for the numerical experiments of this study.

### 4.3 Multilevel Control Variates

It has been shown that MLMC is more cost effective than MC for many stochastic differential equations [65, 49, 24, 128, 21]. Due to the success of this method, we consider continuing along the path of variance – more precisely MSE – reduction by applying control variates to each  $\mathbb{E}[Y_\ell]$  estimate with the aim of reducing the number of samples  $N_\ell$  required by MLMC or the required work to achieve a desired MSE  $\varepsilon^2$ . In detail, given the multilevel correction variable  $Y_\ell$ , as defined in (4.6), and a correlated variable  $Z_\ell$ , a general MLCV correction variable is defined as

$$W_\ell = Y_\ell - \theta_\ell(Z_\ell - \mathbb{E}[Z_\ell]), \quad (4.13)$$

for  $\ell = 1, \dots, L$ , and  $W_0 = Y_0$ . If  $\mathbb{E}[Z_\ell]$  is known, it follows that  $\mathbb{E}[W_\ell] = \mathbb{E}[Y_\ell]$ , and  $\theta_\ell$  can be set such that the MSE of the estimator of  $\mathbb{E}[W_\ell]$ ,  $\hat{W}_\ell$ , is optimized. However, it is often the case that  $\mathbb{E}[Z_\ell]$  is unknown and thus must be estimated. In the following subsections we will discuss our formulation of the control variate  $Z_\ell$  and the approximation of  $\mathbb{E}[Z_\ell]$ . The discussion will include the decomposition of the MSE of  $\hat{W}_\ell$  with respect to  $\mathbb{E}[Y_\ell]$  and the optimal number of samples needed to reach a desired sampling error.

#### 4.3.1 $Z_\ell$ Formulation

We consider a formulation of  $Z_\ell$  based on a low-rank approximation of  $Y_\ell$  in (4.6). Specifically,

$$Z_\ell = Q_\ell^{ID} - Q_{\ell-1}, \quad (4.14)$$

where  $Q_{\ell-1}$  is as defined in (4.6), and  $Q_{\ell}^{ID}$  is a low-rank approximation of  $Q_{\ell}$  based on samples of  $Q_{\ell-1}$  and a pre-determined basis for level  $\ell$  solution described next. Let  $Q_{\ell}$  depend on a solution-dependent, vector-valued quantity  $\mathbf{q}_{\ell}(\boldsymbol{\xi}) \in \mathbb{R}^{m_{\ell}}$ ; that is,

$$Q_{\ell} = Q(\mathbf{q}_{\ell}(\boldsymbol{\xi})).$$

For instance, the (spatial) mean heat flux along a boundary,  $Q_{\ell}$ , can be computed based on the values of the heat flux,  $\mathbf{q}_{\ell}$ , on the boundary grid nodes.

To construct  $Q_{\ell}^{ID}$ , we assume the approximation to  $\mathbf{q}$  on level  $\ell$ ,  $\mathbf{q}_{\ell}$ , admits an accurate representation in a reduced basis of (small) cardinality  $r \ll m_{\ell}$ . Inspired by [91, 142, 28, 27], such a low-rank approximation of  $\mathbf{q}_{\ell}$  may be achieved by identifying a reduced basis for  $\mathbf{q}_{\ell-1}$  along with an approximation rule representing an arbitrary sample of  $\mathbf{q}_{\ell-1}$  in that basis. In details, some  $N_{\ell} \geq m_{\ell}$  realizations of  $\mathbf{q}_{\ell-1}$  corresponding to random samples  $\{\boldsymbol{\xi}^{(i)}\}_{i=1}^{N_{\ell}}$ , are generated and organized in an  $m_{\ell-1} \times N_{\ell}$  *coarse grid data matrix*

$$\mathbf{U}_{\ell-1} := \left[ \mathbf{q}_{\ell-1}(\boldsymbol{\xi}^{(1)}) \quad \mathbf{q}_{\ell-1}(\boldsymbol{\xi}^{(2)}) \quad \cdots \quad \mathbf{q}_{\ell-1}(\boldsymbol{\xi}^{(N_{\ell})}) \right].$$

We note that, here,  $N_{\ell}$  is different from the sample size  $N_{\ell}$  in MLMC and will be defined more precisely in Section 4.3.6.1. To find a reduced basis for  $\mathbf{q}_{\ell-1}$ , we consider a rank  $r$  factorization of  $\mathbf{U}_{\ell-1}$  using a subset of its columns. While a number of tools are available for this purpose [33, 80, 57, 30], we employ the so-called interpolative decomposition (ID) of  $\mathbf{U}_{\ell-1}$  [54, 23, 57],

$$\mathbf{U}_{\ell-1} \approx \mathbf{U}_{\ell-1}^c \mathbf{C}_{\ell-1}, \quad (4.15)$$

where the  $m_{\ell-1} \times r$  *column skeleton matrix*

$$\mathbf{U}_{\ell-1}^c = \left[ \mathbf{q}_{\ell-1}(\boldsymbol{\xi}^{(i_1)}) \quad \mathbf{q}_{\ell-1}(\boldsymbol{\xi}^{(i_2)}) \quad \cdots \quad \mathbf{q}_{\ell-1}(\boldsymbol{\xi}^{(i_r)}) \right]$$

consists of  $r$  columns of  $\mathbf{U}_{\ell-1}$  identified via pivoted rank-revealing QR factorization of  $\mathbf{U}_{\ell-1}$ , and  $\mathbf{C}_{\ell-1}$  is an  $r \times N_{\ell}$  *coefficient matrix* as specified in Appendix A. Stated differently, (4.15) gives a rank  $r$  approximation of  $\mathbf{q}_{\ell-1}(\boldsymbol{\xi}^{(i)})$  in a *reduced basis*  $\mathbf{U}_{\ell-1}^c$  consisting of  $r$  realizations of  $\mathbf{q}_{\ell-1}$ . For

$\mathbf{q}_{\ell-1}(\boldsymbol{\xi})$  evaluated at an arbitrary  $\boldsymbol{\xi}$ , a rank  $r$  representation of the form

$$\mathbf{q}_{\ell-1}(\boldsymbol{\xi}) \approx \mathbf{U}_{\ell-1}^c \mathbf{c}_{\ell-1}(\boldsymbol{\xi}) \quad (4.16)$$

can be generated by computing the  $r$ -vector of coefficients  $\mathbf{c}_{\ell-1}(\boldsymbol{\xi})$  via least squares approximation. Specifically,

$$\mathbf{c}_{\ell-1}(\boldsymbol{\xi}) = \arg \min_{\hat{\mathbf{c}} \in \mathbb{R}^r} \|\mathbf{U}_{\ell-1}^c \hat{\mathbf{c}} - \mathbf{q}_{\ell-1}(\boldsymbol{\xi})\|_2, \quad (4.17)$$

which can be computed by via SVD or QR decomposition. Further discussion related to the complexity of solving for the coefficients,  $\mathbf{c}_{\ell-1}(\boldsymbol{\xi})$ , can be found in Section 4.3.7.

**Remark 4.3.1.** *It is worthwhile highlighting that, in practice,  $r$  is not known a priori and has to be chosen such that the approximation in (4.15) achieves a desired accuracy. Additionally, to reveal the rank  $r$  of  $\mathbf{U}_{\ell-1}$  from (4.15),  $m_{\ell-1}$  must satisfy the condition  $m_{\ell-1} \geq r$ , as  $r \leq \min\{m_{\ell-1}, N_\ell\}$ .*

**Remark 4.3.2.** *The rank  $r$  may vary from one discretization level to another; however, for the interest of a simpler notation, we suppress the dependence of  $r$  on  $\ell$ .*

From this process, we identify samples  $\{\boldsymbol{\xi}^{(i_k)}\}_{k=1}^r$  using which the corresponding *fine grid* reduced basis and approximation can be determined. Specifically, following [91] and using the coarse grid coefficient vector  $\mathbf{c}_{\ell-1}$  in (4.16), we now define the rank  $r$  approximation of  $\mathbf{q}_\ell$  as

$$\mathbf{q}_\ell^{ID}(\boldsymbol{\xi}) = \mathbf{U}_\ell^c \mathbf{c}_{\ell-1}(\boldsymbol{\xi}), \quad (4.18)$$

where the reduced basis

$$\mathbf{U}_\ell^c = \begin{bmatrix} \mathbf{q}_\ell(\boldsymbol{\xi}^{(i_1)}) & \mathbf{q}_\ell(\boldsymbol{\xi}^{(i_2)}) & \cdots & \mathbf{q}_\ell(\boldsymbol{\xi}^{(i_r)}) \end{bmatrix}$$

is the  $m_\ell \times N_\ell$  fine grid counterpart of  $\mathbf{U}_{\ell-1}^c$ . Finally, (4.14) is fully specified by setting

$$Q_\ell^{ID} = Q(\mathbf{q}_\ell^{ID}(\boldsymbol{\xi})).$$

For a discussion on the convergence of the coarse-fine approximation  $\mathbf{q}_\ell^{ID}$  in (4.18), we refer the interested reader to Sections 3 and 4 of [91], where an identical construction is presented.

### 4.3.2 $\mathbb{E}[Z_\ell]$ Estimation

The numerical construction of  $Z_\ell$  in (4.14) does not lead to an analytic value for  $\mathbb{E}[Z_\ell]$ , as desired in standard control variate approaches. However, approximations to  $\mathbb{E}[Z_\ell]$ , denoted hereafter by  $\bar{Z}_\ell$ , may be alternatively used as long as they are not expensive to generate [120, 34, 98]. While there are multiple techniques to estimate  $\mathbb{E}[Z_\ell]$ , an independent MC simulation is used here. In detail, let  $N'_\ell$  be the number of MC samples collected from level  $\ell - 1$  to estimate  $\mathbb{E}[Z_\ell]$ . We define the MC estimator of  $\mathbb{E}[Z_\ell]$  to be

$$\bar{Z}_\ell = \frac{1}{N'_\ell} \sum_{i=1}^{N'_\ell} Z_\ell^{(i)} = \frac{1}{N'_\ell} \sum_{i=1}^{N'_\ell} \left( Q_\ell^{ID(i)} - Q_{\ell-1}^{(i)} \right), \quad (4.19)$$

where  $Q_\ell^{ID(i)}$  is generated following the procedure of Section 4.3.1. An advantage to this MC method is that the estimate  $\bar{Z}_\ell$  only relies on additional level  $\ell - 1$  samples without any need to draw additional level  $\ell$  samples. Thus, given the reduced bases  $\mathbf{U}_{\ell-1}^c$  and  $\mathbf{U}_\ell^c$ , the primary cost of finding  $\bar{Z}_\ell$  is running the level  $\ell - 1$  model to obtain  $N'_\ell$  coarse grid samples, which is given by

$$\mathcal{C}(\bar{Z}_\ell) = N'_\ell \mathcal{C}(Q_\ell^{ID}) + N'_\ell \mathcal{C}(Q_{\ell-1}) \approx N'_\ell \mathcal{C}(Q_{\ell-1}).$$

A more detailed quantification of  $\mathcal{C}(Q_\ell^{ID})$  will be provided in Section 4.3.7. Next, we focus on the MSE reduction achieved with this method of control variates, and subsequently, the estimation of  $\tilde{N}_\ell$  and  $N'_\ell$ .

### 4.3.3 MSE Reduction via $W_\ell$

Now that we have defined  $Z_\ell$  and  $\bar{Z}_\ell$  we can update (4.13) to get

$$W_\ell = Y_\ell - \theta_\ell(Z_\ell - \bar{Z}_\ell), \quad (4.20)$$

for  $\ell = 1, \dots, L$ , and  $W_0 = Y_0$ . Given a fixed (independently computed)  $\bar{Z}_\ell$ ,  $\hat{W}_\ell$  is a biased estimator of  $\mathbb{E}[Y_\ell]$  with the bias  $\theta_\ell \delta_\ell$ , where  $\delta_\ell = \bar{Z}_\ell - \mathbb{E}[Z_\ell]$ . However,  $\delta_\ell$  may be thought of as a random variable with mean zero and variance  $\mathbb{V}[Z_\ell]/N'_\ell$ . Define  $\tilde{N}_\ell$  to be the number of samples of  $(Y_\ell, Z_\ell)$  from (4.20) and recall  $N'_\ell$  is the number of samples used to calculate  $\bar{Z}_\ell$ . Based on the work in



[98],  $\theta_\ell$  can be selected such that the MSE of  $\hat{W}_\ell$  – also averaged over possible realizations of  $\delta_\ell$  – is minimized.

By doing so, it follows that, see [98],

$$\text{MSE}(\hat{W}_\ell, \mathbb{E}[Y_\ell]) = \text{MSE}(\hat{Y}_\ell, \mathbb{E}[Y_\ell]) \left[ 1 - \rho_\ell^2 \left( \frac{1}{1 + \tilde{N}_\ell/N'_\ell} \right) \right], \quad (4.21)$$

where

$$\rho_\ell^2 = \frac{(\text{cov}(Y_\ell, Z_\ell))^2}{\mathbb{V}[Z_\ell]\mathbb{V}[Y_\ell]} \quad (4.22)$$

and

$$\theta_\ell^* = \frac{\text{cov}(Y_\ell, Z_\ell)}{\mathbb{V}[Z_\ell]} \left( \frac{1}{1 + \tilde{N}_\ell/N'_\ell} \right). \quad (4.23)$$

Since  $0 \leq \rho_\ell^2 \leq 1$ , (4.21) and (4.22) indicate that as the correlation between  $Y_\ell$  and  $Z_\ell$  increases,  $\rho_\ell^2$  approaches one, and thus the *MSE reduction factor*  $\text{MSE}(\hat{W}_\ell, \mathbb{E}[Y_\ell])/\text{MSE}(\hat{Y}_\ell, \mathbb{E}[Y_\ell])$  goes to zero.

Updating  $\hat{W}_\ell$  to account for the optimal MSE, the MLCV correction estimator is given by

$$\hat{W}_\ell = \frac{1}{\tilde{N}_\ell} \sum_{i=1}^{\tilde{N}_\ell} \left( Y_\ell^{(i)} - \theta_\ell^* (Z_\ell^{(i)} - \bar{Z}_\ell) \right), \quad (4.24)$$

for  $\ell = 1, \dots, L$ , and  $\hat{W}_0 = \hat{Y}_0$ . Then the full MLCV estimator is defined as

$$\hat{Q}_L^{MLCV} = \sum_{\ell=0}^L \hat{W}_\ell.$$

In order to complete the task of calculating the number of samples  $\tilde{N}_\ell$  per level to optimize the cost, we must find the MSE of this MLCV estimator first. It can be shown that

$$\text{MSE}(\hat{Q}_L^{MLCV}, \mathbb{E}[Q]) = \text{MSE}(\hat{Q}_L^{MLCV}, \mathbb{E}[Q_L]) + (\mathbb{E}[Q_L - Q])^2,$$

where the MSE's are also averaged over the realizations of  $\delta_\ell$ . Expanding the sampling error term,

we find

$$\begin{aligned}
\text{MSE}(\hat{Q}_L^{MLCV}, \mathbb{E}[Q_L]) &= \mathbb{E} \left[ \left( \hat{Q}_L^{MLCV} - \mathbb{E}[Q_L] \right)^2 \right] \\
&= \mathbb{E} \left[ \left( \hat{W}_0 - \mathbb{E}[Y_0] \right)^2 \right] + \cdots + \mathbb{E} \left[ \left( \hat{W}_L - \mathbb{E}[Y_L] \right)^2 \right] \\
&= \sum_{\ell=0}^L \text{MSE}(\hat{W}_\ell, \mathbb{E}[Y_\ell]) \\
&= \sum_{\ell=0}^L \text{MSE}(\hat{Y}_\ell, \mathbb{E}[Y_\ell]) \left[ 1 - \rho_\ell^2 \left( \frac{1}{1 + \tilde{N}_\ell/N'_\ell} \right) \right],
\end{aligned}$$

where the last line is determined from (4.21). Thus the MSE of  $\hat{Q}_L^{MLCV}$  is given by

$$\text{MSE}(\hat{Q}_L^{MLCV}, \mathbb{E}[Q]) = \sum_{\ell=0}^L \frac{1}{\tilde{N}_\ell} \mathbb{V}[Y_\ell] \left[ 1 - \rho_\ell^2 \left( \frac{1}{1 + \tilde{N}_\ell/N'_\ell} \right) \right] + (\mathbb{E}[Q_L] - Q)^2, \quad (4.25)$$

where we used the relation  $\text{MSE}(\hat{Y}_\ell, \mathbb{E}[Y_\ell]) = \mathbb{V}[Y_\ell]/\tilde{N}_\ell$ . Since the discretization error term in (4.25) is the same as in (4.8), the MLMC convergence theory, as discussed in Section 4.2.3, can be applied to the simulations of  $W_\ell$  in MLCV in order to guarantee a bound on the MSE.

#### 4.3.4 Number of Samples $\tilde{N}_\ell$ for $\hat{W}_\ell$

In MLMC the number of samples needed for each level is estimated in order to attain a bound on the sampling error component of the MSE, see (4.12). In this formulation of MLCV, we must not only account for the sample size  $\tilde{N}_\ell$ , but also for  $N'_\ell$ . The sample size  $\tilde{N}_\ell$  is the required number of samples of  $(Y_\ell, Z_\ell)$  to achieve a sampling error bound of  $\varepsilon^2/2$ , while the value  $N'_\ell$  is the number of samples used to find  $\bar{Z}_\ell$ .

In practice, we are able to determine the ratio  $\tilde{N}_\ell/N'_\ell$  from a number of pilot runs, as discussed in Section 4.3.6, and update the MSE in (4.25) accordingly. Applying the same method to derive (4.12) and fixing the ratio  $\tilde{N}_\ell/N'_\ell$ , the optimal number of samples  $\tilde{N}_\ell$  for  $\hat{W}_\ell$  can be determined from (4.25) and is given by

$$\tilde{N}_\ell = \frac{2}{\varepsilon^2} \left[ \sum_{k=0}^L \left( \mathbb{V}[Y_k] \left[ 1 - \rho_k^2 \left( \frac{1}{1 + \tilde{N}_k/N'_k} \right) \right] \mathcal{C}(W_k) \right)^{1/2} \right] \sqrt{\frac{\mathbb{V}[Y_\ell] \left[ 1 - \rho_\ell^2 \left( \frac{1}{1 + \tilde{N}_\ell/N'_\ell} \right) \right]}{\mathcal{C}(W_\ell)}}, \quad (4.26)$$

where  $\rho_k^2$  is also estimated from the pilot simulations. We note that in (4.26), we take

$$\mathcal{C}(W_\ell) = \mathcal{C}_\ell.$$

This cost accounts for the samples of  $Y_\ell$  and  $Z_\ell$ . We consider the data acquisition for  $\bar{Z}_\ell$  to be offline, thus the related cost is not included in the estimation of  $\tilde{N}_\ell$ ; however, it is included in the total cost of implementing MLCV, as will be discussed in Section 4.3.7.

#### 4.3.5 Number of Samples $N'_\ell$ for $\bar{Z}_\ell$

As can be observed from (4.25) or (4.26), the ratio  $\tilde{N}_\ell/N'_\ell$  plays an important role in the MSE reduction achieved by  $\hat{W}_\ell$  and the number of samples  $\tilde{N}_\ell$  required by MLCV on each level. The smaller  $\tilde{N}_\ell/N'_\ell$ , the larger the MSE reduction and, hence, the smaller the estimated  $\tilde{N}_\ell$  from (4.26). However, a small ratio  $\tilde{N}_\ell/N'_\ell$  requires a large number  $N'_\ell$  of independent samples to compute  $\bar{Z}_\ell$  and, hence, a large overall computational cost. Therefore, it is necessary to determine an *optimal* value for  $\tilde{N}_\ell/N'_\ell$ .

In [98], the authors discuss the use of the so-called *generalized* MSE to find the optimal number of samples to minimize the product of the cost and  $\text{MSE}(\hat{W}_\ell, \mathbb{E}[Y_\ell])$  from (4.21). Let  $\zeta = \mathcal{C}(Q_{\ell-1})/\mathcal{C}_\ell$  denote the cost per sample of  $Z_\ell$  (the cost of a level  $\ell - 1$  sample) divided by the cost per sample of  $Y_\ell$ , where  $\mathcal{C}_\ell$  is given in (4.11). Following [98] and by minimizing the product of the total cost required on each level – given by  $\zeta N'_\ell + (1 + \zeta)\tilde{N}_\ell$  – and  $\text{MSE}(\hat{W}_\ell, \mathbb{E}[Y_\ell])$ , the optimal ratio  $\tilde{N}_\ell/N'_\ell$  is such that

$$N'_\ell = \max(0, s_1 - 1) \tilde{N}_\ell \tag{4.27}$$

with

$$s_1 = \left[ \frac{\rho_\ell^2}{\zeta(1 - \rho_\ell^2)} \right]^{1/2}.$$

If  $N'_\ell = 0$  this implies we do not implement MLCV on level  $\ell$ , instead, we let  $W_\ell = Y_\ell$ .

The potential issue with using (4.27) is that for  $\rho_\ell^2 \approx 1$ , the value of  $N'_\ell$  can get quite large. In some tests this may occur, and depending on the desired implementation, the number of samples

may alternatively be determined by

$$N'_\ell = s_2 \tilde{N}_\ell, \quad (4.28)$$

where the integer  $s_2 > 1$  is preselected. In the numerical experiments of this work, we choose the minimum  $N'_\ell$  from (4.27) and (4.28), i.e.,

$$N'_\ell = \min(s_2, \max(0, s_1 - 1)) \tilde{N}_\ell, \quad (4.29)$$

with  $s_2 = 10$ . This selected value of  $s_2 = 10$  will insure more than 90% of  $\rho_\ell^2$  is maintained in the MSE reduction factor (see (4.21)), when  $s_2$  is used to evaluate (4.29). For values  $s_2 > 10$ , the benefit of maintaining a slightly larger percentage of  $\rho_\ell^2$  is more outweighed by the computational cost of generating additional solution realizations.

#### 4.3.6 MLCV Implementation Details and Algorithms

In this section we discuss the implementation of MLCV as outlined in Algorithms 2 and 3. The MLCV algorithm can be organized into two different components: the pilot run and the full MLCV run. For both, we will describe the case that  $\ell > 0$ . When  $\ell = 0$ , we simply simulate the level  $\ell = 0$  model and take a sample average to calculate  $\hat{W}_0 = \hat{Y}_0$ . Additionally, we assume that  $L$  is large enough to meet the desired discretization error. In practice, we check that convergence is achieved by estimating  $\alpha$  and  $\beta$ , as introduced in Section 4.2.3.

##### 4.3.6.1 Pilot Run – Reduced Basis and Sample Size Determination

As in the case of MLMC, the goal of the pilot run is to find estimates of  $\mathbb{E}[Y_\ell]$  and  $\mathbb{V}[Y_\ell]$ , to ensure level  $L$  is fine enough such that the discretization error is bounded by  $\varepsilon^2/2$ , and to determine the sample sizes  $\tilde{N}_\ell$  and  $N'_\ell$ . However, determining  $\tilde{N}_\ell$  and  $N'_\ell$  requires the estimation of  $\rho_\ell^2$  and the derivation of the reduced basis  $\mathbf{U}_\ell^c$  (along with the associated interpolation rules) to construct the control variates  $Z_\ell$ . These are specific to this MLCV approach and will be additionally performed during the pilot run stage summarized in Algorithm 2.

In particular, a small set of  $N_p \geq m_{L-1}$  samples of  $\boldsymbol{\xi}$  is generated and applied to level  $\ell - 1$  and  $\ell$  models to obtain samples of  $Y_\ell$ , as well as estimates of  $\mathbb{E}[Y_\ell]$  and  $\mathbb{V}[Y_\ell]$ . We note that  $N_p$  may vary from one level to the other, provided that some prior knowledge of  $\mathbb{V}[Y_\ell]$  is available. As discussed in Section 4.3.1, a reduced basis  $\mathbf{U}_{\ell-1}^c$  and an associated interpolation rule for the level  $\ell - 1$  solution vector  $\mathbf{q}_{\ell-1}$  is found via the ID of the coarse grid data matrix  $\mathbf{U}_{\ell-1}$ . The level  $\ell$  reduced basis  $\mathbf{U}_\ell^c$  is generated by simulating the level  $\ell$  model at the samples of  $\boldsymbol{\xi}$  identified by the level  $\ell - 1$  ID and used – in conjunction with the same interpolation rule – to acquire samples of  $\mathbf{q}_\ell^{ID}$ , and subsequently  $Q_\ell^{ID}$  and  $Z_\ell$ . These bases are stored and reused in the MLCV runs of Section 4.3.6.2. Following this,  $\rho_\ell^2$  and  $\tilde{N}_\ell/N'_\ell$  are computed from (4.22) and (4.29), respectively. After each level the change in the sample means of  $Y_\ell$  must be determined to verify that the intended discretization accuracy is met. If not, this process is repeated with larger values of  $L$ . If the value of  $L$  is sufficient, the pilot run ends and the values of  $\tilde{N}_\ell$  and  $N'_\ell$  are determined from (4.26) and (4.22), respectively.

**Remark 4.3.3.** *While the variance and covariance are estimated from the pilot run, these values, in practice, may be updated during the main MLCV run described next.*

#### 4.3.6.2 Main MLCV Run

After the total number of samples  $\tilde{N}_\ell$  and  $N'_\ell$  for each level are estimated, the full MLCV run is completed following Algorithm 3. Much like the pilot run,  $\tilde{N}_\ell$  realizations of  $Y_\ell^{(i)}$  are collected from the level  $\ell$  and  $\ell - 1$  models. Using the  $\tilde{N}_\ell$  samples of  $\mathbf{q}_{\ell-1}$ , and the level  $\ell - 1$  reduced basis  $\mathbf{U}_{\ell-1}^c$  from the pilot run, least squares (4.17) is performed on each sample to find the coefficient vectors  $\mathbf{c}_{\ell-1}(\boldsymbol{\xi})$ . Given this interpolation rule and the corresponding level  $\ell$  basis  $\mathbf{U}_\ell^c$ , the samples of  $\mathbf{q}_\ell^{ID}$  (hence  $Q_\ell^{ID}$ ) and subsequently  $Z_\ell^{(i)}$  are generated using (4.18) and (4.14), respectively. Following these,  $\bar{Z}_\ell$  is computed from (4.19) using  $N'_\ell$  independent samples of  $Z_\ell$  drawn as described above. After calculating  $\theta_\ell^*$  in (4.23), we have all the components of the MLCV estimator, and use (4.24) to find  $\hat{W}_\ell$  and lastly  $\hat{Q}_L^{MLCV}$ .

**Algorithm 2:** MLCV Reduced Basis Identification and Sample Size Determination

```

for  $\ell = 0, \dots, L$  do
  if  $\ell = 0$  then
     $\{\xi^{(i)}\}_{i=1}^{N_p} \rightarrow \{Q_\ell^{(i)}\}_{i=1}^{N_p} \rightarrow \{Y_\ell^{(i)}\}_{i=1}^{N_p}$ 
  else
     $\{\xi^{(i)}\}_{i=1}^{N_p} \rightarrow \{\mathbf{q}_{\ell-1}^{(i)}\}_{i=1}^{N_p}, \{\mathbf{q}_\ell^{(i)}\}_{i=1}^{N_p} \rightarrow \{Q_{\ell-1}^{(i)}\}_{i=1}^{N_p}, \{Q_\ell^{(i)}\}_{i=1}^{N_p} \rightarrow \{Y_\ell^{(i)}\}_{i=1}^{N_p}$ 
     $\mathbf{U}_{\ell-1} = [\mathbf{q}_{\ell-1}^{(1)} \cdots \mathbf{q}_{\ell-1}^{(N_p)}] \xrightarrow{\text{ID via (4.15)}} \mathbf{U}_{\ell-1}^c \mathbf{C}_{\ell-1}, \{\xi^{(i_k)}\}_{i=1}^r$ 
     $\{\xi^{(i_k)}\}_{i=1}^r \rightarrow \mathbf{U}_\ell^c; \mathbf{U}_\ell^{ID} = \mathbf{U}_\ell^c \mathbf{C}_{\ell-1} \rightarrow \{Q_\ell^{ID(i)}\}_{i=1}^{N_p} \rightarrow \{Z_\ell^{(i)}\}_{i=1}^{N_p}$ 
    Calculate  $\rho_\ell^2$  and  $\tilde{N}_\ell/N'_\ell$ , respectively, via (4.22) and (4.29)
  end
  Estimate  $\mathbb{V}[Y_\ell]$ 
end
for  $\ell = 0, \dots, L$  do
  Calculate  $\tilde{N}_\ell$  using (4.26)
  Calculate  $N'_\ell$  using (4.29)
end

```

**Algorithm 3:** Main MLCV Run

```

for  $\ell = 0, \dots, L$  do
  if  $\ell = 0$  then
     $\{\xi^{(i)}\}_{i=1}^{\tilde{N}_\ell} \rightarrow \{Q_\ell^{(i)}\}_{i=1}^{\tilde{N}_\ell} \rightarrow \{Y_\ell^{(i)}\}_{i=1}^{\tilde{N}_\ell} \rightarrow \hat{W}_\ell = \hat{Y}_\ell$ 
  else
    %  $\bar{Z}_\ell$  calculation:
     $\{\xi^{(i)}\}_{i=1}^{N'_\ell} \rightarrow \{\mathbf{q}_{\ell-1}^{(i)}\}_{i=1}^{N'_\ell} \rightarrow \{Q_{\ell-1}^{(i)}\}_{i=1}^{N'_\ell}$ 
     $\mathbf{c}_{\ell-1}^{(i)} = \min_{\hat{\mathbf{c}} \in \mathbb{R}^r} \|\mathbf{U}_{\ell-1}^c \hat{\mathbf{c}} - \mathbf{q}_{\ell-1}^{(i)}\|$  for  $i = 1, \dots, N'_\ell$ 
     $\mathbf{q}_\ell^{ID(i)} = \mathbf{U}_\ell^c \mathbf{c}_{\ell-1}^{(i)}$  for  $i = 1, \dots, N'_\ell \rightarrow \{Q_\ell^{ID(i)}\}_{i=1}^{N'_\ell} \rightarrow \{Z_\ell^{(i)}\}_{i=1}^{N'_\ell} \rightarrow \bar{Z}_\ell$ 
    %  $Z_\ell$  sampling:
     $\{\xi^{(i)}\}_{i=1}^{\tilde{N}_\ell}$  (independent of  $\{\xi^{(i)}\}_{i=1}^{N'_\ell}$  above)  $\rightarrow \{\mathbf{q}_{\ell-1}^{(i)}\}_{i=1}^{\tilde{N}_\ell}, \{\mathbf{q}_\ell^{(i)}\}_{i=1}^{\tilde{N}_\ell} \rightarrow$ 
     $\{Q_{\ell-1}^{(i)}\}_{i=1}^{\tilde{N}_\ell}, \{Q_\ell^{(i)}\}_{i=1}^{\tilde{N}_\ell} \rightarrow \{Y_\ell^{(i)}\}_{i=1}^{\tilde{N}_\ell}$ 
     $\mathbf{c}_{\ell-1}^{(i)} = \min_{\hat{\mathbf{c}} \in \mathbb{R}^r} \|\mathbf{U}_{\ell-1}^c \hat{\mathbf{c}} - \mathbf{q}_{\ell-1}^{(i)}\|$  for  $i = 1, \dots, \tilde{N}_\ell$ 
     $\mathbf{q}_\ell^{ID(i)} = \mathbf{U}_\ell^c \mathbf{c}_{\ell-1}^{(i)}$  for  $i = 1, \dots, \tilde{N}_\ell \rightarrow \{Q_\ell^{ID(i)}\}_{i=1}^{\tilde{N}_\ell} \rightarrow \{Z_\ell^{(i)}\}_{i=1}^{\tilde{N}_\ell}$ 
    %  $W_\ell$  sampling:
     $\{W_\ell^{(i)}\}_{i=1}^{\tilde{N}_\ell}$  from (4.24)  $\rightarrow \hat{W}_\ell$ 
  end
end
Set  $\hat{Q}_L^{MLCV} = \sum_{\ell=0}^L \hat{W}_\ell$ 

```

### 4.3.7 MLCV Cost Breakdown

After determining the total number of samples required to optimize the cost with the MSE bound as a constraint, the total cost of implementing MLMC is given in (4.10). By modifying MLMC through the use of the control variate  $Z_\ell$ , we are able to achieve further reduction in the MSE of the estimator  $\hat{W}_\ell$ , as long as  $Z_\ell$  is sufficiently correlated with  $Y_\ell$ . In doing so, the sample size for  $Y_\ell$  and thus  $Z_\ell$  needed to meet the sampling error bound, while optimizing the total cost, is reduced. However, sampling  $Z_\ell$  and estimating its expectation, as required in (4.13), lead to additional computational cost that we quantify next.

*Cost of pilot run* Recall that the construction of  $Z_\ell$  is based on the identification of the reduced basis  $\mathbf{U}_{\ell-1}^c$  (using which  $\mathbf{U}_\ell^c$  is generated) and the associated least squares coefficients  $\mathbf{c}_{\ell-1}$  computed from (4.17). Following Algorithm 2, the former is based on the ID of the coarse data matrix  $\mathbf{U}_{\ell-1}$  consisting of  $N_p$  pilot samples of  $\mathbf{q}_{\ell-1} \in \mathbb{R}^{m_{\ell-1}}$ . From [84], the cost to implement a rank  $r$  ID on  $\mathbf{U}_{\ell-1}$  is  $\mathcal{O}(m_{\ell-1}\tilde{N}_\ell n_r)$ , where  $n_r < \tilde{N}_\ell$ . This means that the cost per sample of  $Z_\ell$  is  $\mathcal{O}(m_{\ell-1}n_r)$ . In many scenarios of practical interest, the QoI  $Q$  depends on a  $\mathbf{q}$  consisting of a considerably smaller set of solution degrees of freedom, e.g., average outflow temperature or lift/drag on an airfoil, as opposed to those of the entire solution. In such cases,  $m_{\ell-1} \ll M_{\ell-1}$  and we may ignore the cost of performing the ID of  $\mathbf{U}_{\ell-1}$ . Assuming that (i) the factorization to solve (4.17) is formed only once, e.g, using a reduced QR decomposition with cost  $\mathcal{O}(r^2 m_{\ell-1})$ , (ii)  $r \ll m_\ell$ , and (iii)  $\tilde{N}_\ell \gg 1$ , the cost of solving for the coefficients  $\mathbf{c}_{\ell-1}$  for each sample and applying them to  $\mathbf{U}_\ell^c$  to obtain  $\mathbf{q}_\ell^{ID}$  is  $\mathcal{O}(r m_\ell)$ , which again may be ignored when  $m_\ell \ll M_\ell$ . From the  $N_p$  pilot samples,  $r$  of them form the reduced bases  $\mathbf{U}_{\ell-1}^c$  and  $\mathbf{U}_\ell^c$  the rest will be recycled in the main MLCV run, assuming  $N_p \leq \tilde{N}_\ell$ . Therefore, the net cost of performing the pilot run on each level is the same as the cost of generating  $r$  samples of level  $\ell$  and  $\ell - 1$  solutions to form the reduced bases  $\mathbf{U}_{\ell-1}^c$  and  $\mathbf{U}_\ell^c$ . This gives rise to the total cost of  $\sum_{l=1}^L r (\mathcal{C}(Q_{\ell-1}) + \mathcal{C}(Q_\ell))$  for the pilot run.

*Cost of main MLCV run* Sampling  $Z_\ell$  in (4.20) entails using the  $\tilde{N}_\ell$  samples of  $\mathbf{q}_{\ell-1}$  drawn for  $Y_\ell$  and performing least squares regression to generate the approximation (4.16). As discussed

previously, these least squares solves have negligible cost when  $m_\ell \ll M_\ell$ . Therefore, there is no additional cost for sampling  $Z_\ell$  in (4.20) relative to its MLMC counterpart. With a similar argument, sampling  $Z_\ell$  in (4.20) to compute  $\bar{Z}_\ell$  requires only generating  $N'_\ell$  independent samples of  $\mathbf{q}_{\ell-1}$ . Therefore, the cost of the main run is  $\tilde{N}_0 \mathcal{C}(Q_0) + \sum_{\ell=1}^L \tilde{N}_\ell (\mathcal{C}(Q_{\ell-1}) + \mathcal{C}(Q_\ell)) + \sum_{\ell=1}^L N'_\ell \mathcal{C}(Q_{\ell-1})$ .

*Total cost* Adding the cost of pilot and main MLCV runs, we arrive at the total cost of MLCV given by

$$\mathcal{C}(\hat{Q}_L^{MLCV}) = \tilde{N}_0 \mathcal{C}(Q_0) + \sum_{\ell=1}^L (\tilde{N}_\ell + r) (\mathcal{C}(Q_{\ell-1}) + \mathcal{C}(Q_\ell)) + \sum_{\ell=1}^L N'_\ell \mathcal{C}(Q_{\ell-1}). \quad (4.30)$$

#### 4.3.8 Discussion on Use of MLCV

Before proceeding to the numerical experiments and results, we first discuss the types of problems that benefit from the application of MLCV. When deciding to apply a MC based method to solve a high-dimensional problem in uncertainty quantification, the practitioner must consider the QoI's variance – more specifically coefficient of variation (COV) – as well as the cost of simulating the model on a fine enough discretization. For problems with small COV and cheap-to-evaluate models, it can be the case that standard MC will perform as well, or may outperform MLMC. On the other hand, the practitioner will likely decide to use MLMC over MC if the QoI shows a large COV and the solver exhibits a fast cost growth between refined grids ( $\mathcal{C}_\ell \gg M_\ell$ ). Since MLCV is a multilevel method, it benefits from the same traits as MLMC: a large COV and a fast cost growth. However, unlike MLMC, MLCV relies on a low-rank representation of the solution, in the sense defined in Section 4.3.1, thus making the approach most effective in outperforming MLMC on problems exhibiting small solution ranks on all the levels. When this occurs we find that  $Q_\ell^{ID}$  approximates the fine grid data  $Q_\ell$  accurately, and that the value for  $\rho_\ell^2$  is much closer to one, thus resulting in an improved MSE for  $\hat{W}_\ell$ .

A final comment must be made regarding the total cost of MLCV. In the case that significant MSE reduction for  $\hat{W}_\ell$  does occur, we are not necessarily guaranteed cost reduction. In (4.30) the number of samples needed on levels  $\ell > 0$  is always larger than  $r$ . In MLMC, this constraint is not present (see (4.10)). For scenarios where the data exhibits a small variance or small values of



$\varepsilon$  are provided as the desired tolerance, the resulting number of samples on the finer levels may be comparable to  $r$ . In such cases, MLCV may not outperform MLMC. This suggests MLCV is more beneficial for smaller selections of  $\varepsilon$ .

#### 4.4 Numerical Results

In this section we consider two test cases to compare the resulting error and cost estimates for MLMC and MLCV. In the first case we consider the generalized eigenvalue problem for finding the frequency of the first natural mode of vibration of an L-shaped elastic structure. For the second case we consider a thermally driven cavity flow with stochastic boundary temperatures previously studied in [82, 115, 100, 60].

Following each case, we present several results when comparing MLCV and MLMC for both tests. In particular, we consider the convergence of  $\rho_\ell^2$  in (4.22), and subsequently the MSE at each level, which controls the sampling error of MLCV. For values of  $\rho_\ell^2$  close to one, a significant reduction in the MSE of  $\hat{W}_\ell$  can be observed, and thus a reduction in the number  $\tilde{N}_\ell$  of required samples per level. We further provide results on the total cost and relative error of MLCV and MLMC. The total cost, as a function of  $\varepsilon$ , is determined from (4.26), (4.29), and (4.30). For the eigenvalue problem, the cost is defined to be  $\mathcal{C}_\ell = M_\ell$ , while the thermally driven flow problem cost is approximated by an average CPU time. The final result is the relative error in estimating the mean of QoI – to be specified – on all discretization levels. The errors are generated based on a reference solution  $\hat{Q}_L^{\text{ref}}$  obtained via sparse polynomial chaos expansion [29, 60] on the finest available discretization for each model. Specifically,

$$e_\ell^{\text{MLMC}} = \frac{\left| \sum_{k=0}^{\ell} \hat{Y}_k - \hat{Q}_L^{\text{ref}} \right|}{\left| \hat{Q}_L^{\text{ref}} \right|} \quad (4.31)$$

and

$$e_\ell^{\text{MLCV}} = \frac{\left| \sum_{k=0}^{\ell} \hat{W}_k - \hat{Q}_L^{\text{ref}} \right|}{\left| \hat{Q}_L^{\text{ref}} \right|}, \quad (4.32)$$

for  $\ell = 1, \dots, L$ .

#### 4.4.1 Test 1: Natural Frequency of a 2D Linear Elasticity Problem

For the first test, we seek to find the expected value of the smallest vibration frequency,  $\omega$ , of the linear elasticity problem on an L-shape domain  $\mathcal{D}$  as shown in Figure 4.1 (a). We consider zero Dirichlet boundary conditions on the top and right sides of the domain along  $y$  and  $x$  directions, respectively. The Young's modulus of the medium, given by the log-normal random field

$$E(\mathbf{x}, \boldsymbol{\xi}) = \bar{E} + \exp(G(\mathbf{x}, \boldsymbol{\xi})), \quad \mathbf{x} \in \mathcal{D},$$

is the source of uncertainty in this test. Here,  $G(\mathbf{x}, \boldsymbol{\xi})$  is a Gaussian random field represented by the Karhunen-Loève expansion

$$G(\mathbf{x}, \boldsymbol{\xi}) = \sum_{i=1}^d \sqrt{\lambda_i} \phi_i(\mathbf{x}) \xi_i,$$

where the random variables  $\xi_i$  are independent standard Gaussian and  $\lambda_i$  are the  $d$  largest eigenvalues with corresponding eigenfunctions  $\phi_i(\mathbf{x})$  determined by the Gaussian covariance function

$$\mathcal{K}(\mathbf{x}_1, \mathbf{x}_2) = \sigma^2 \exp\left(-\frac{\|\mathbf{x}_1 - \mathbf{x}_2\|_2^2}{l}\right).$$

In our experiments, we set  $\bar{E} = 0.1$ ,  $d = 36$ ,  $\sigma^2 = 0.33$ , and  $l = 0.33$ , and assume the medium has a unit density.

To compute  $\omega$ , we consider the solution of the random eigenvalue problem

$$\mathbf{K}(\boldsymbol{\xi}) \mathbf{q}(\boldsymbol{\xi}) = \omega^2(\boldsymbol{\xi}) \mathbf{M} \mathbf{q}(\boldsymbol{\xi}), \quad (4.33)$$

where  $\mathbf{K}(\boldsymbol{\xi})$  and  $\mathbf{M}$  are, respectively, the stiffness and mass matrices associated with finite element discretization of the linear elasticity equations on uniform triangular meshes of varying size. Additionally,  $\mathbf{q}(\boldsymbol{\xi})$  is the natural vibration mode (eigenvector) corresponding to the eigenvalue  $\omega^2(\boldsymbol{\xi})$  given by

$$\omega^2(\boldsymbol{\xi}) = \frac{\mathbf{q}(\boldsymbol{\xi})^T \mathbf{K}(\boldsymbol{\xi}) \mathbf{q}(\boldsymbol{\xi})}{\mathbf{q}(\boldsymbol{\xi})^T \mathbf{M} \mathbf{q}(\boldsymbol{\xi})}. \quad (4.34)$$

Here, QoI is the smallest frequency  $\omega$  satisfying (4.33) and is a functional of  $\mathbf{q}(\boldsymbol{\xi})$ , i.e.  $Q = Q(\mathbf{q}(\boldsymbol{\xi}))$ , through (4.34). With the given uncertainty, the COV of  $Q$  is approximately 15%.

The finite element discretization is done using FENiCS [79] with five different meshes of varying size, as displayed in Figure 4.1 (b)-(f). The number of degrees of freedom of the meshes are  $M_0 = 36$ ,  $M_1 = 106$ ,  $M_2 = 212$ ,  $M_3 = 856$ , and  $M_4 = 3296$ . The resulting eigenvector  $\mathbf{q}_\ell$ , from each level  $\ell$  simulation, has  $M_\ell$  degrees of freedom. After eigenvectors  $\mathbf{q}_\ell^{ID}$  are found, the corresponding (saved) stiffness and mass matrices,  $\mathbf{K}$  and  $\mathbf{M}$ , are used in (4.34) to generate  $Q_\ell^{ID}$  and the control variates  $Z_\ell$ .

For the application of MLMC and MLCV we consider two cases to test the relative performance of these methods when altering the difference in the number of degrees of freedom between adjacent levels. In the first case, *A*, we apply MLMC and MLCV to all five levels. In the second case, *B*, we consider only using three of the five meshes from levels  $\ell = 0, 2, 4$ . This is worth testing, as both methods rely on the differences between QoIs on adjacent levels. The results between *A* and *B* will indicate which method can adjust to the larger gap in the number of degrees of freedom between levels for this problem.

To select the rank on each level, we consider values of the correlation  $\rho_\ell^2$  (see (4.22)) for different values of  $r$ , based on pilot run data. Figure 4.2 demonstrates how the values of  $\rho_\ell^2$  change with rank  $r$  for case *A* (left) and case *B* (right). In order to have  $\rho_\ell^2 > 0.90$  we use  $r = 10$  for all levels, with the exception of  $r = 15$  on the finest level,  $\ell = 4$ , in case *A*.

#### 4.4.1.1 Results of Test 1

For the results presented in this section, the “o” markers denote case *A* data and the “x” markers denote case *B* data. The first set of results are related to the MLMC convergence guarantee in Section 4.2.3. Figure 4.3 (left) shows the sample mean of  $Q_\ell$ ,  $Y_\ell$  and  $W_\ell$  as a function of the number of degrees of freedom on each level,  $M_\ell$ . We note that the sample mean of  $Q_\ell$  remains relatively constant, while those of  $Y_\ell$  and  $W_\ell$  decay at the same rate. The data displayed allows us to

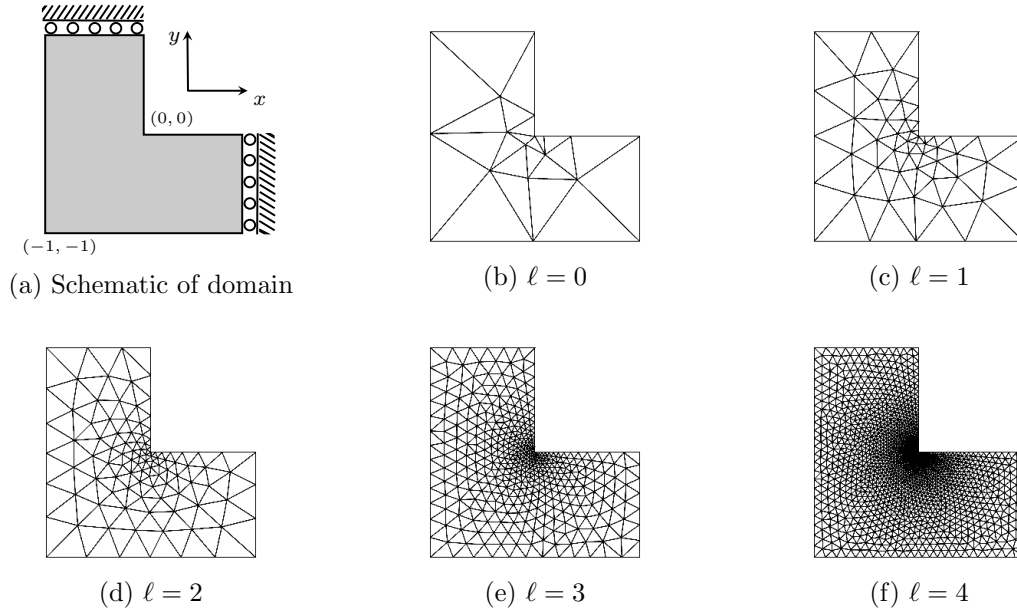


Figure 4.1: (a) Schematic of the L-shaped domain. (b)-(f) Meshes on the five different levels with increasing resolution.

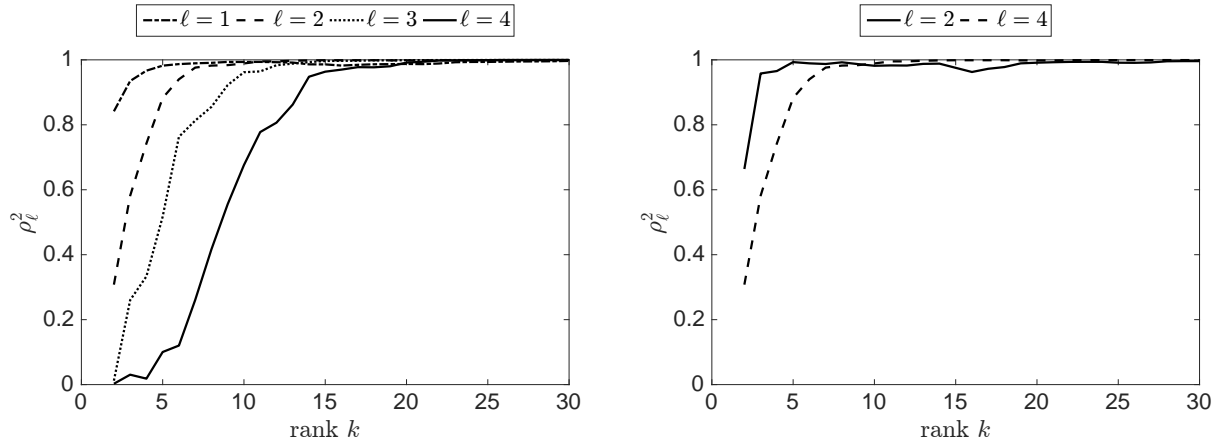


Figure 4.2: Change of the correlation  $\rho_\ell^2$  in (4.22) as a function of the ID rank  $r$  for (left) case  $A$  and (right) case  $B$ .

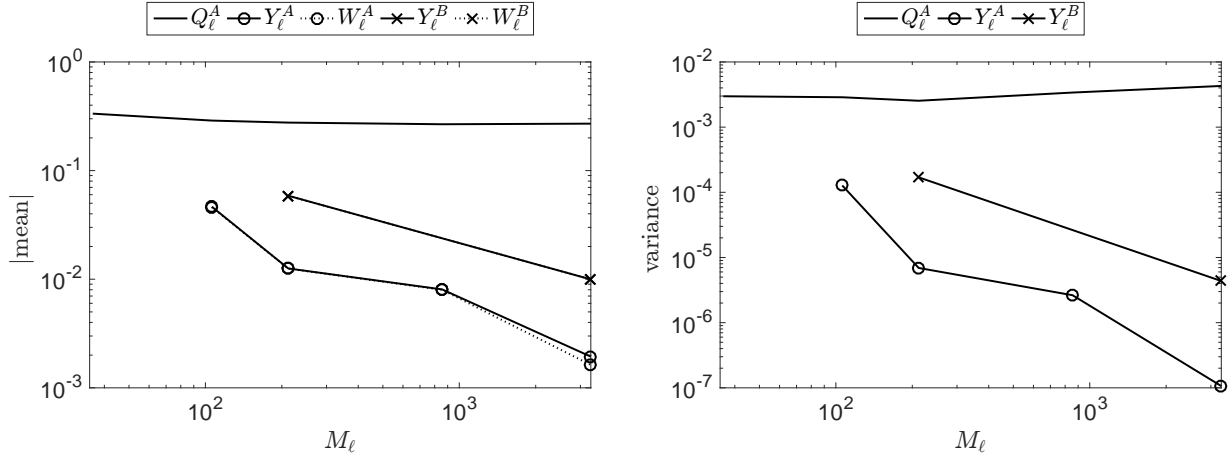


Figure 4.3: (Left) Convergence of the sample mean of  $Y_\ell$  and  $W_\ell$ , while that of  $Q_\ell$  remains relatively constant. (Right) Convergence of the sample variance of  $Y_\ell$  in comparison to that of  $Q_\ell$ . Superscripts  $A$  and  $B$  refer to the case  $A$  and  $B$  of the L-shaped domain problem, respectively.

estimate the values of  $\alpha$ . Based on the slope, we have  $|\mathbb{E}[Q_\ell - Q_{\ell-1}]| \approx M_\ell^{-0.92}$  and  $|\mathbb{E}[Q_\ell - Q_{\ell-1}]| \approx M_\ell^{-0.64}$  for case  $A$  and  $B$ , respectively. Thus it follows that  $\alpha_A \approx 0.92$  and  $\alpha_B \approx 0.64$ . Figure 4.3 (right) shows the sample variance of  $Q_\ell$  and  $Y_\ell$  as a function of the degrees of freedom on each level,  $M_\ell$ . The data displayed allows us to estimate the values of  $\beta$ . From Figure 4.3 (right) we have  $\mathbb{V}[Q_\ell - Q_{\ell-1}] \approx M_\ell^{-1.7}$  and  $\mathbb{V}[Q_\ell - Q_{\ell-1}] \approx M_\ell^{-1.3}$ , for case  $A$  and  $B$ , respectively. This implies that  $\beta_A \approx 1.7$  and  $\beta_B \approx 1.3$ . The cost of the FE solver, as well as the cost to compute  $\omega_1^2$ , indicate that  $\gamma \approx 1$ , i.e.  $\mathcal{C}_\ell \lesssim M_\ell$ , suggesting that for case  $A$  and  $B$  we may consider a tolerance of  $\varepsilon \sim \mathcal{O}(0.0005)$ . We remark that, while the value of  $\alpha_B$  indicates the discretization error cannot meet the specified tolerance, we argue that this is an estimation of  $\mathbb{E}[Q_L - Q]$ , and that this approximation of  $\alpha$  is better obtained from test  $A$  results.

Next we consider the MSE reduction in Figure 4.4. Figure 4.4 (left) shows the increasing values of  $\rho_\ell^2$  as a function of  $\ell$ , as well as the decreasing values of the MSE reduction factor. For both cases  $A$  and  $B$ ,  $\rho_\ell^2 > 0.95$  for  $\ell > 0$ . This is ideal, as this reduction in the MSE means that fewer samples are required to obtain the MSE tolerance. Figure 4.4 (right) displays the number of samples required on each level for both methods when using  $\varepsilon = 0.0005$ . Note that the number of

samples for MLCV includes those used to calculate  $\bar{Z}_\ell$ . For case *A* the number of samples required for MLCV is slightly smaller than that of MLMC. A more significant reduction in the number of samples needed for MLCV is observed for case *B*. This suggests that MLCV is more adaptable when large gaps are added between  $M_\ell$  on adjacent levels.

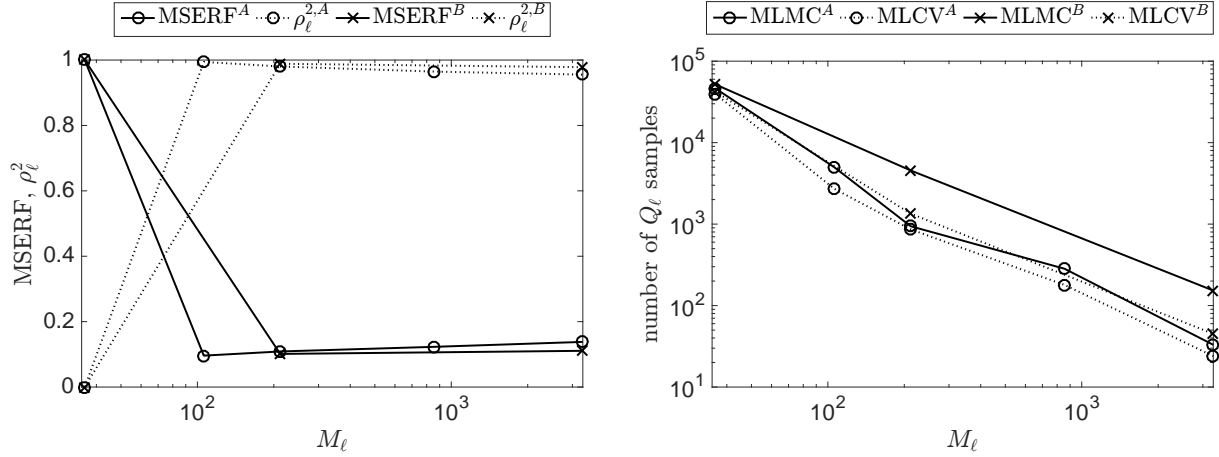


Figure 4.4: (Left) MSE reduction factor (MSERF) between  $\hat{W}_\ell$  and  $\hat{Y}_\ell$ , as well as the value of  $\rho_\ell^2$  on each level. (Right) Total number of samples needed to achieve the sampling error  $\varepsilon = 0.0005$  for MLMC and MLCV. MLCV sample size includes the number required to determine  $\bar{Z}_\ell$ . Superscripts *A* and *B* refer to the case *A* and *B* of the L-shaped domain problem, respectively.

To compare the costs for these two methods, we turn to Table 4.1, which displays the cost estimates for several cases as a function of prescribed  $\varepsilon$  based on (4.10) and (4.30), where  $\mathcal{C}_\ell = M_\ell$ . For comparison, the cost of MC on the finest level is also reported. For small values of  $\varepsilon$ , an improvement in cost is observed for MLCV. When requiring a tolerance of  $\varepsilon = 0.0005$  we observe a cost ratio of 75% for case *A* and 58% for case *B*. Comparing the total costs of these two cases, we see that the MLCV method for case *B* outperforms case *A*. For MLMC, it is the opposite. By adding larger gaps between levels, the performance of MLMC decays, while that of MLCV improves. When increasing the tolerance, there is a reduction in the cost gain of MLCV, as fewer samples are needed to maintain the MSE tolerance. The cost of the  $r$  samples for the basis is diminishing the success

$\varepsilon$	Levels	Cost MC	Cost MLMC	Cost MLCV	$\frac{\text{Cost MLCV}}{\text{Cost MLMC}}$
0.0005	0, 1, 2, 3, 4	8.7e7	2.8e6	2.1e6	0.75
0.0005	0, 2, 4	6.6e7	3.3e6	1.9e6	0.58
0.001	0, 1, 2, 3	2.2e7	6.5e5	5.0e5	0.77
0.003	0, 1, 2, 3	2.4e6	7.3e4	7.1e4	0.97
0.005	0, 1, 2, 3	8.7e5	2.7e4	3.7e4	1.40

Table 4.1: Cost of MC, MLMC, and MLCV for Test 1 problem. The difference in cost of MC for  $\varepsilon = 0.0005$  is due to the small number of samples used.

of the MLCV method. When  $\varepsilon > 0.003$ , the cost of performing MLCV is greater than the cost of performing MLMC.

These next results compare the estimated MSEs and relative errors of the methods. Figure 4.5 (left) displays the estimated MSE for both methods. For both case  $A$  and  $B$ , the MSE estimates for  $\hat{Y}_\ell$  and  $\hat{W}_\ell$  are on the same order, which is expected. The slight difference is due to the number of terms in the sampling error. Since case  $A$  has more levels, most of the sampling error terms should be smaller than those of case  $B$ , as the total sampling error for both test cases must attain the same tolerance of  $\varepsilon^2/2$ .

To see how this new MLCV method performs against MLMC in terms of relative error, we compare the values computed in (4.31) and (4.32). Figure 4.5 (right) displays the convergence of the relative error for both methods as the levels are refined. We note that, due to the cost of simulating the fine model, these results are based on an average of 10 runs of MLMC and MLCV, where the data in each run is not completely independent from data in other runs. This result indicates that, experimentally, both methods converge on the same order. However the two methods do not cost the same to derive this result. As stated earlier, using the cost and the total number of samples on each level, it is found that to achieve a sampling error of  $\varepsilon = 0.0005$ , MLCV only requires about 75% and 58% of the computational cost of MLMC for case  $A$  and  $B$ , respectively.

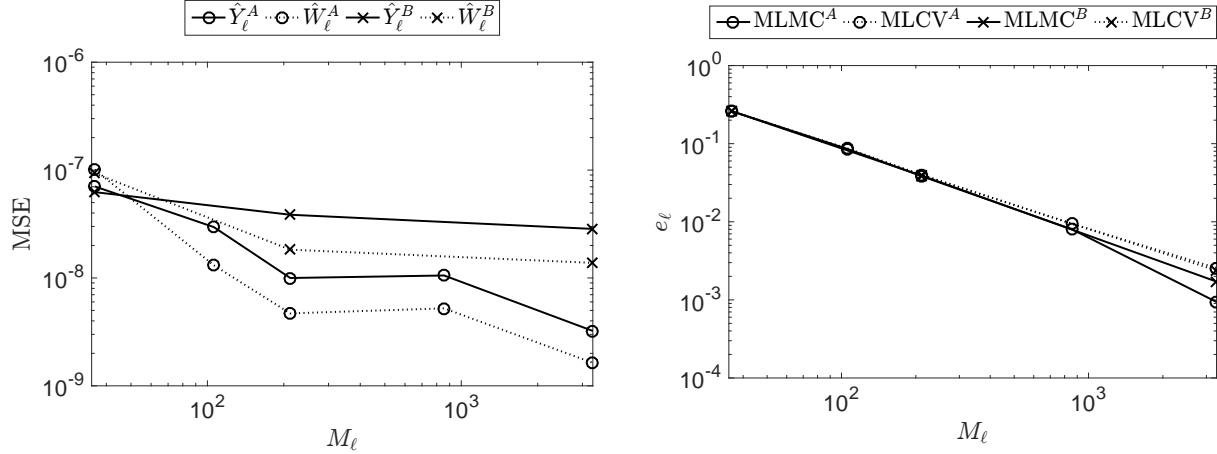


Figure 4.5: (Left) MSE for  $\hat{Y}_\ell$  and  $\hat{W}_\ell$ . (Right) Comparison of the relative errors (4.31) and (4.32) in estimating the mean QoI at each level for MLMC and MLCV. Superscripts  $A$  and  $B$  refer to the case  $A$  and  $B$  of the L-shaped domain problem, respectively.

#### 4.4.2 Test 2: Thermally Driven Cavity Flow

For the second implementation of MLCV, we consider a thermally driven flow problem in a square domain as described in [108, 81] and illustrated in Figure 4.6. The left vertical wall has a random temperature  $T_h$  with mean  $\bar{T}_h$ , while the right vertical wall, referred to as the cold wall, has a spatially varying stochastic temperature  $T_c < T_h$  with constant mean  $\bar{T}_c$ . Both top and bottom walls are assumed to be adiabatic. The reference temperature and the reference temperature difference are defined as  $T_{ref} = \bar{T}_c$  and  $\Delta T_{ref} = T_h - \bar{T}_c$ , respectively. Under small temperature difference assumption, i.e., Boussinesq approximation, the normalized governing equations are given by, [81],

$$\begin{aligned} \frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} &= -\nabla p + \frac{\text{Pr}}{\sqrt{\text{Ra}}} \nabla^2 \mathbf{u} + \text{Pr} \Theta \mathbf{e}_y, \\ \nabla \cdot \mathbf{u} &= 0, \\ \frac{\partial \Theta}{\partial t} + \nabla \cdot (\mathbf{u} \Theta) &= \frac{1}{\sqrt{\text{Ra}}} \nabla^2 \Theta, \end{aligned} \tag{4.35}$$

where  $\mathbf{e}_y$  is the unit vector  $(0, 1)$ ,  $\mathbf{u} = (u, v)$  is velocity vector field,  $\Theta = (T - T_{ref})/\Delta T_{ref}$  is normalized temperature,  $p$  is pressure, and  $t$  is time. Zero velocity boundary conditions on all walls (in both directions) are assumed. For more details on the normalization of the variables



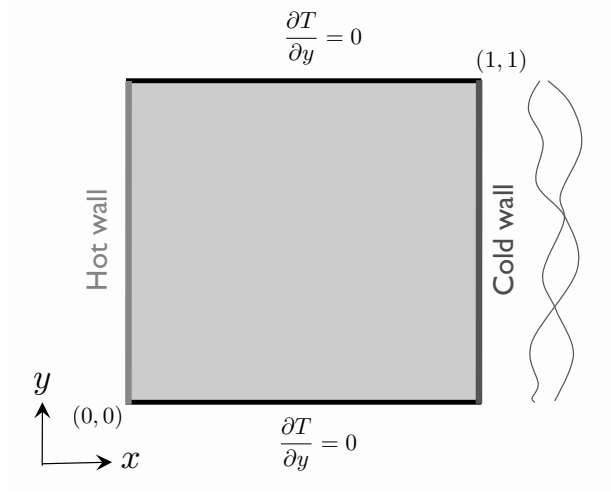


Figure 4.6: Schematic of Test 2, the thermally driven flow problem with random temperature along the hot (west) wall and spatially varying stochastic temperature along the cold (east) wall.

in (4.35), we refer the interested reader to [108, 81]. Prandtl and Rayleigh numbers are defined, respectively, as  $\text{Pr} = \nu/\alpha$  and  $\text{Ra} = g\tau\Delta T_{ref}L^3/(\nu\alpha)$ . Specifically,  $L$  is the length of the cavity,  $g$  is gravitational acceleration,  $\nu$  is kinematic viscosity,  $\alpha$  is thermal diffusivity, and the coefficient of thermal expansion is given by  $\tau$ . In this example, we set  $g = 10$ ,  $L = 1$ ,  $\tau = 0.5$ , and  $\text{Pr} = 0.71$ . We use finite volume for the discretization of (4.35). The QoI  $Q$  is the spatial average of steady-state heat flux along the hot wall.

The uncertain quantities include the initial temperatures on the cold and hot wall ( $T_c$  and  $T_h$ , respectively), as well as the viscosity parameter  $\nu$ . The cold wall temperature,  $T_c$ , is given by a Karhunen-Loève-type expansion with 50 terms,

$$T_c(y, \boldsymbol{\xi}) = \bar{T}_c + \sigma \sum_{i=1}^{50} \sqrt{\lambda_i} \phi_i(y) \xi_i, \quad (4.36)$$

where  $\lambda_i$  and  $\phi_i(y)$  are the eigenvalues and eigenfunctions of the exponential correlation function

$$\mathcal{K}(y_1, y_2) = \sigma^2 \exp\left(-\frac{|y_1 - y_2|}{l}\right), \quad (4.37)$$

with  $\sigma = 2$  and  $l = 0.15$ . In (4.36), we set  $\bar{T}_c = 100$  and assume the random variables  $\xi_i \sim U[-1, 1]$ ,  $i = 1, \dots, 50$ , are independently and uniformly distributed over  $[-1, 1]$ . Additionally, we assume

$T_h \sim U[105, 109]$  and  $\nu \sim U[0.004, 0.01]$ . This brings the total number of random inputs to  $d = 52$ . We note that the COV for the QoI of each simulation with this setup is about 25%.

For the numerical setup of this problem, we consider four nested uniform grids with the sizes of  $16 \times 16$ ,  $32 \times 32$ ,  $64 \times 64$ , and  $128 \times 128$ . This implies we have  $M_\ell = s^\ell M_0$ , with  $s = 4$ . The resulting vector of heat flux values along the hot wall,  $\mathbf{q}$ , is of length  $m_\ell = \sqrt{M_\ell}$ . As stated earlier, since  $m_\ell \ll M_\ell$ , the added cost of performing ID and least squares is discounted.

To select the rank on each level, we consider values of the correlation  $\rho_\ell^2$  (see (4.22)) for different values of  $r$ , based on pilot run data. Figure 4.7 demonstrates how the values of  $\rho_\ell^2$  change with rank  $r$ . In order to have  $\rho_\ell^2 > 0.90$  we use  $r = 10$  for all levels.

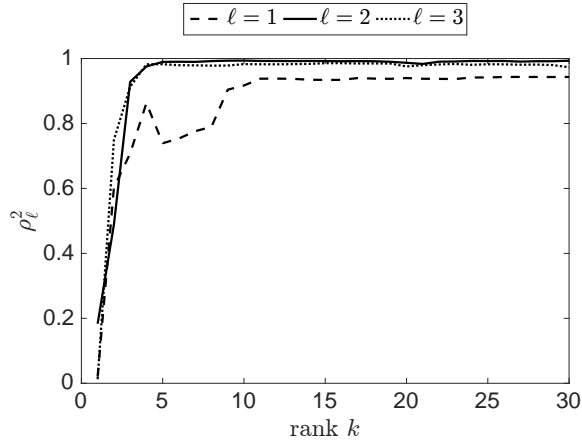


Figure 4.7: Change of the correlation  $\rho_\ell^2$  in (4.22) as a function of the ID rank  $r$  for Test 2.

#### 4.4.2.1 Results of Test 2

We first consider the results corresponding to the MLMC convergence guarantee in Section 4.2.3. Figure 4.8 (left) shows the sample mean of  $Q_\ell$ ,  $Y_\ell$  and  $W_\ell$  as a function of the degrees of freedom on each level,  $M_\ell$ . We note that the sample mean of  $Q_\ell$  remains relatively constant, while those of  $Y_\ell$  and  $W_\ell$  decay at the same rate. The data displayed allows us to estimate the values

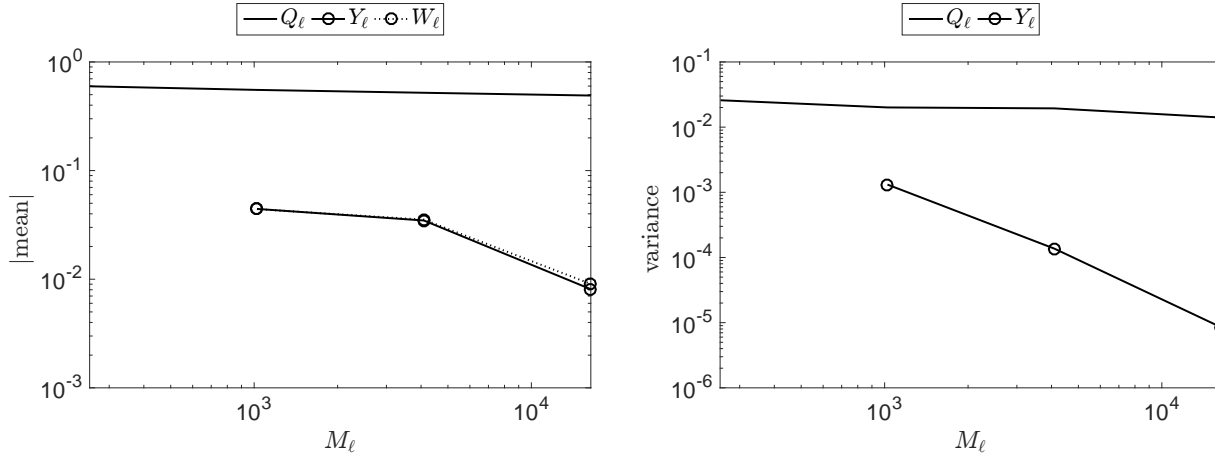


Figure 4.8: (Left) Convergence of the sample mean of  $Y_\ell$  and  $W_\ell$ , while that of  $Q_\ell$  remains relatively constant. (Right) Convergence of the sample variance of  $Y_\ell$  in comparison to that of  $Q_\ell$ .

of  $\alpha$ . Since the decay of data  $Y_\ell$ ,  $\ell = 1, 2$ , and  $3$  does not show a clear convergence, i.e.,  $\ell = 0$  discretization is *too coarse*, we use the data based on  $Y_2$  and  $Y_3$  to approximate  $\alpha$ . By doing so, we estimate  $\mathbb{E}[Q_\ell - Q_{\ell-1}] \approx M_\ell^{-1.0}$ , indicating that  $\alpha \approx 1.0$ . We note this estimate was then confirmed by using data obtained from the  $256 \times 256$  mesh. Figure 4.8 (right) shows the sample variance of  $Q_\ell$  and  $Y_\ell$  as a function of the degrees of freedom on each level,  $M_\ell$ . The data displayed allows us to estimate the values of  $\beta$ . We estimate  $\mathbb{V}[Q_\ell - Q_{\ell-1}] \approx M_\ell^{-1.8}$ , indicating that  $\beta \approx 1.8$ . By using the average CPU time to determine the cost of the solver, we find that  $\gamma \approx 1.65$ . This implies that the minimum MSE bound we can have is  $\varepsilon \sim \mathcal{O}(0.00005)$  when using  $M_L = 128^2$ . Due to cost constraints, we select  $\varepsilon = 0.001$  for the numerical experiments.

Next we consider the MSE reduction as well as the sample size reduction in Figure 4.9. Figure 4.9 (left) shows the values of  $\rho_\ell^2$ , where  $\rho_\ell^2 > 0.85$  for levels  $1, 2$ , and  $3$ , as well as the decay of the MSE reduction factor. As the MSE reduction factor approaches zero, the number of samples MLCV requires on each level decays. Figure 4.9 (right) displays the number of samples required on each level for both methods when using  $\varepsilon = 0.001$ . It is clear that fewer samples are needed on

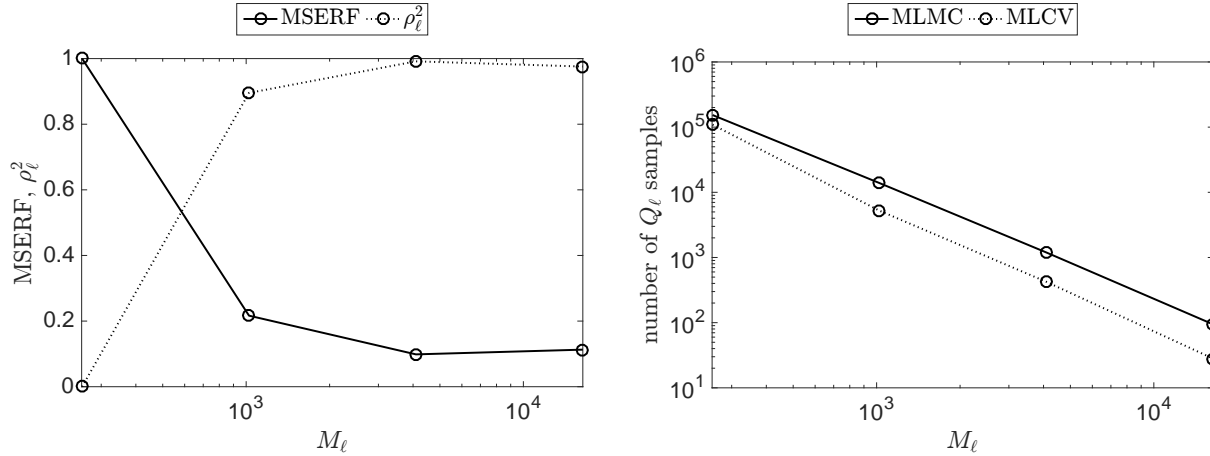


Figure 4.9: (Left) MSE reduction factor (MSERF) between  $\hat{W}_\ell$  and  $\hat{Y}_\ell$ , as well as the value of  $\rho_\ell^2$  on each level. (Right) Total number of samples needed to attain the sampling error for MLMC and MLCV when  $\varepsilon = 0.001$ . MLCV sample size includes the number required to determine  $\bar{Z}_\ell$ .

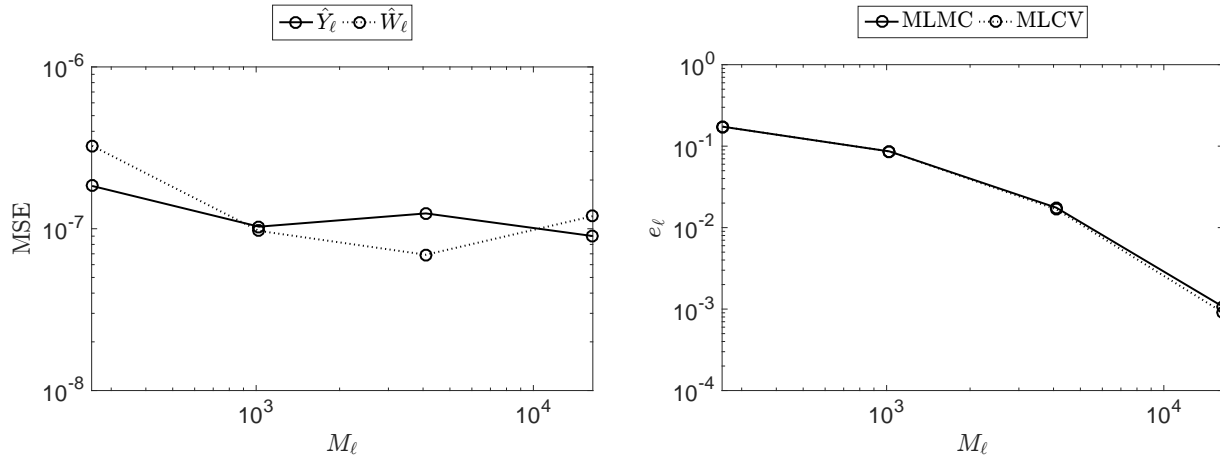
each level for MLCV than for MLMC. And, as finer levels are approached, the gap between these two values increases. Because of this, we will see a significant cost payoff.

To compare the cost of performing MLMC and MLCV for this application, we consider the cost estimates displayed in Table 4.2, as determined by (4.10) and (4.30). Due to high simulation costs, we consider only four levels. We see that for  $\varepsilon \leq 0.001$  MLCV halves the cost of MLMC. As the value of  $\varepsilon$  is increased, the cost difference of the two methods is smaller. For  $\varepsilon \geq 0.005$ , we can observe that MLCV has a larger computational cost than MLMC. The cost of the  $r$  samples for the basis is diminishing the success of the MLCV method.

The final set of results compare the estimated MSEs and relative errors of the two methods. Figure 4.10 (left) displays the estimated MSE for both methods. We see that the MSE estimates for  $\hat{Y}_\ell$  and  $\hat{W}_\ell$  are on the same order. To determine the relative accuracy of the two methods, we next compute the relative errors (4.31) and (4.32). Figure 4.10 (right) displays the convergence of the relative error for both methods as the levels are refined. We note that, due to the cost of simulating the fine model, these results are based on an average of 10 runs of MLMC and MLCV, where the

$\varepsilon$	Levels	Cost MC	Cost MLMC	Cost MLCV	$\frac{\text{Cost MLCV}}{\text{Cost MLMC}}$
0.00005	0, 1, 2, 3	2.2e9	3.7e7	1.7e7	0.45
0.0005	0, 1, 2, 3	2.2e7	3.7e5	1.7e5	0.45
0.001	0, 1, 2, 3	5.6e6	9.3e4	4.4e4	0.48
0.003	0, 1, 2, 3	6.2e5	1.0e4	6.8e3	0.65
0.005	0, 1, 2, 3	2.2e5	3.8e3	4.0e3	1.1

Table 4.2: Cost of MC, MLMC, and MLCV for Test 2 problem.

Figure 4.10: (Left) MSE for  $\hat{Y}_\ell$  and  $\hat{W}_\ell$ . (Right) Comparison of the relative errors (4.31) and (4.32) in estimating the mean of QoI at each level for MLMC and MLCV.

data in each run is not completely independent from the other runs. As expected, we observe that as the mesh is refined for both methods, the relative errors improve. The notable result is that the relative errors for both methods remain essentially the same, while, in fact, MLCV only requires 48% of the computational cost of MLMC.

#### 4.4.3 Analysis

The preceding results show a cost improvement over MLMC when applying this MLCV method. In addition, other observations regarding these results are important to highlight. In Test

1, the difference between the number of degrees of freedom on adjacent levels has a greater negative impact on the MLMC results, than those of MLCV. When using MLCV, the process does not need to be broken up into as many levels with small changes in the degrees of freedom as in MLMC, as MLCV does not rely solely on the difference between QoIs on adjacent levels. This method is especially beneficial in cases where there is not sufficient flexibility in defining multiple coarse spatial discretizations, or when the solution converges slowly over the refinements.

The cost ratio between MLCV and MLMC is also significantly different when comparing Test 1 and Test 2. While the  $\rho_\ell^2$  results for both methods are favorable, the faster cost growth between levels in Test 2 results in greater cost savings. More precisely, in Test 1 we had  $\gamma = 1$  and in Test 2 we had  $\gamma = 1.65$ . The payoff of placing more computational burden on the coarse level data than that of the fine level is more exaggerated for larger values of  $\gamma$ . This also suggests that MLCV will be even more advantageous over MLMC for three-dimensional (in physical space) problems, where  $M_\ell$  grows faster.

## 4.5 Conclusion

In this paper we have developed a non-intrusive, multilevel Monte Carlo method to approach uncertainty quantification problems with a large number of random inputs. The mathematical framework of this method relies heavily on the decomposition of the coarse level QoI data into a low-rank representation, from which a relatively accurate fine grid estimator can be calculated. Data with fast decaying singular values, combined with the use of control variates to minimize the MSE of the new variable results in this new MLCV method, with a total cost that is smaller than or comparable with that of MLMC.

Both MLCV and MLMC methods are applied to a generalized eigenvalue problem associated with a linear elasticity model as well as a thermally driven cavity flow problem in a square domain. While the data in both tests lends itself to approximations with different ranks, and costs to determine the QoIs, MLCV was found to outperform MLMC. To reach a fixed MSE tolerance, MLCV was determined to require fewer samples, and thus a cheaper implementation than that of

MLMC.

## Chapter 5

### A Bi-Fidelity Approximation for Linear Bayesian Inference

#### 5.1 Introduction

Characterizing and incorporating uncertainties when simulating physical phenomena is essential for improving model-based predictions. These uncertainties typically correspond to physical parameters that describe the system, e.g., variable temperature or velocities, or other quantities that are not precisely known. The task of forward uncertainty quantification (UQ) is concerned with treating these uncertainties as stochastic model parameters, where the probability distribution describing these parameters is assumed to be known. Typically this information is inferred from knowledge of the underlying physics or experimental data, and termed the prior probability distribution. Resulting numerical solutions are thus stochastic by design, and the aim of forward UQ is estimating the statistics of the associated quantity of interest determined from these solutions. Inverse UQ, on the other hand, uses numerical solutions to infer the statistics on the input parameters. When observational data is available, predictions may be further improved by incorporating this information via Bayesian inference. This information allows for the update of the prior probability distribution describing the uncertainty, where this updated probability distribution is referred to as the posterior.

For problems in which this uncertainty, or more appropriately termed, unknown, is of high-dimension, performing computational methods in Bayesian inference can be infeasible. Notably, with an increase in the number of unknowns, the more computationally demanding the forward problem becomes. Typically the number of unknowns corresponds to the degrees of freedom in



the numerical discretization of the governing equations and the unknown parameters. A specific challenge, and the primary interest in this chapter, is estimating the posterior covariance of the unknown parameter, given a set of observations. For a linear parameter-to-observable map, described with Gaussian prior and noise, the posterior mean and covariance are estimated via maximum a posteriori (MAP) estimates. The posterior covariance is given by

$$\mathbf{\Gamma}_{\text{post}} = \left( \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A} + \mathbf{\Gamma}_{\text{prior}}^{-1} \right)^{-1}, \quad (5.1)$$

where  $\mathbf{A} \in \mathbb{R}^{m \times N}$  is the linear parameters-to-observables map that maps the unknown parameters to the observations with  $m \ll N$ ,  $\mathbf{\Gamma}_{\text{prior}} \in \mathbb{R}^{N \times N}$  is the prior covariance with assumed structure describing the unknown parameter, and  $\mathbf{\Gamma}_{\text{noise}} \in \mathbb{R}^{m \times m}$  is the covariance describing the noise of the observational data. The posterior covariance in (5.1) may, naïvely, be formed by a matrix inversion with a computational complexity of  $\mathcal{O}(N^3)$ . For large values of  $N$ , both the formation and inversion of this matrix become prohibitive, and model reduction methods are necessary to estimate the posterior covariance.

Over the years several works have focused on model reduction methods for improved cost and scalability of estimating the posterior covariance. Typical methods include utilizing surrogate models, multilevel methods, or dimension reduction techniques, as these methods rely on model reduction via formulation of alternative cheaper models, estimation of the posterior covariance in a multilevel setting, and exploitation of low-rank structure of  $\mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A}$  and structure of  $\mathbf{\Gamma}_{\text{prior}}$  (in (5.1)) to avoid direct inversion to form the posterior covariance, respectively.

Several methods have been studied as surrogates in Bayesian inference. While many require an extensive offline stage to develop the surrogate, a significant advantage of these methods is quick online simulations. This is of great value when exploring the parameter space. Examples of such work include the use of reduced basis models for the forward solver, and using these models in place of the parameter-to-observables map  $\mathbf{A}$  with the addition of exploiting the low-rank structure of  $\mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A}$  [78]. Other work that develops surrogate models the Bayesian inverse can be found in [85, 86, 69]; however, these methods do not discuss further methods for cost reduction of forming

the posterior covariance outside of reducing the cost of the forward solver. A drawback to these methods is that they may require the formation of a new model, instead of treating the model as a black-box solver.

Multilevel methods incorporate solvers utilizing lower resolution discretization schemes, sometimes referred to as coarse grid solvers. As the coarse grid model may not provide an accurate solution, it will still provide some pertinent information regarding the observations, and captures the dynamics of the system. For instance, work has been done to solve the posterior covariance via multigrid methods [13, 4]. Work in [7] incorporates discretization error into the noise assumed to be in the parameter-to-observables mapping. A multilevel method is introduced in [31], where preconditioned Markov chain Monte Carlo employs solutions of a coarse grid solver to aid in the acceptance-rejection process, resulting in an improvement in acceptance rate of solutions to the fine grid solver.

Current state-of-the-art methods rely on transforming the inversion problem in (5.1) into a rank  $r$  eigenvalue decomposition. In particular, the authors of [40, 18] rewrite the posterior covariance as

$$\mathbf{\Gamma}_{\text{post}} = \mathbf{\Gamma}_{\text{prior}}^{1/2} \left( \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A} \mathbf{\Gamma}_{\text{prior}}^{1/2} + \mathbf{I} \right)^{-1} \mathbf{\Gamma}_{\text{prior}}^{1/2}, \quad (5.2)$$

and seek to calculate the first  $r$  eigenvalues  $\mathbf{\Lambda}_r$  and eigenvectors  $\mathbf{V}_r$  of the prior preconditioned data misfit of the Hessian  $\mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A} \mathbf{\Gamma}_{\text{prior}}^{1/2}$ . Then, applying the Sherman-Morrison-Woodbury formula, the  $N \times N$  inverse in (5.2) may be approximated as

$$\left( \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A} \mathbf{\Gamma}_{\text{prior}}^{1/2} + \mathbf{I} \right)^{-1} \approx \mathbf{I} - \mathbf{V}_r (\mathbf{I} + \mathbf{\Lambda}_r)^{-1} \mathbf{\Lambda}_r \mathbf{V}_r^T.$$

This method provided in these works has been pivotal, as it reduces the cost from an inversion to a problem requiring forward and adjoint model solves, where the number of such solves relies only on the rank of the system, which is typically significantly smaller than the size of the linear system. Furthermore, the use of matrix-free methods to calculate the eigenvalue decomposition, e.g., the Lanczos algorithm, improves the cost and efficiency of calculating the eigenmodes. An extension to extreme-scale is performed in [19], where the authors employ a randomized Lanczos algorithm

(following ideas of [77, 57]), resulting in improved efficiency and fault tolerance. Other notable work that explores the use of such low-rank approximations is found in [123] that focuses on identifying optimal formulation of this low-rank method, [12] improves the complexity for problems with time stepping, [117] uses a hierarchical matrix approach for improved complexity, and [118] obtains further complexity improvement through their formulation of the prior covariance. However, for high-cost solvers, the cost of such methods may still be exceedingly high, and a research area of interest is how to exploit coarse grid models, in conjunction with the formulation of this low-rank approximations for further cost reduction.

A direction yet to be explored is how to exploit the similarities of low-fidelity and high-fidelity models, or more appropriately, fine and coarse grid resolution models, to obtain the eigenmodes of  $\mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A} \mathbf{\Gamma}_{\text{prior}}^{1/2}$ . For scenarios where there are fewer observations,  $m$ , than unknowns,  $N$ , a low-fidelity version of the prior preconditioned data misfit of the Hessian matrix may be formed from a coarse grid version of  $\mathbf{A}$  and  $\mathbf{\Gamma}_{\text{prior}}$ . Similar ideas may be found in [140]; however, their work relates to kriging and co-kriging methods, where the covariance matrix has a different formulation, and they use column subset selection for the Nyström approximation.

### 5.1.1 Contributions of this work

In this work, we introduce a novel bi-fidelity approach for approximating the posterior covariance, that relies on a lower fidelity model in a black-box framework. More precisely, we leverage low-fidelity model solves in addition to high-fidelity model solves to approximate the low-rank eigenvalue decomposition to the data misfit of the Hessian. Following the low-rank methodologies of [40, 18, 19] the approximation of the posterior covariance can be reduced to forming a rank  $r$  eigenvalue decomposition of the Hessian misfit (as will be discussed in Section 5.3). Inspired by the Nyström approximation, which provides an approximation to the Hessian misfit via a subset of columns as in [134, 111] or via a mapping between the left and right singular vectors as in [42, 141], this approximation considers a similar type of mapping. However, we utilize a lifting approach, where the left singular vectors and singular values calculated from the low-fidelity model

are mapped to an approximation of the right singular vectors of the high-fidelity model, i.e., the eigenvectors of the data misfit of the Hessian. From this, a bi-fidelity approximation to the posterior covariance may be formed with a reduced computational cost.

To show the success of this method, we consider the following outline for this chapter. In Section 5.2 we provide background and context for linear Bayesian inference. In Section 5.3 we outline the low-rank methodologies of [40, 18] for calculating the posterior covariance approximation. In Section 5.4 we present this novel algorithm which incorporates low-fidelity models, as well as theoretical error bound for this approximation. Section 5.4.2 provides a cost analysis of this method versus other state-of-the-art in posterior covariance approximation. Section 5.5 provides numerical examples to compare the accuracy of this method as well as cost with other methods.

## 5.2 Bayesian Inference for Linear Gaussian Problems

We first provide a brief background on Bayesian inference for determining the posterior density. Consider the linear mapping,  $A$ , between two random variables  $X$  and  $Y$ , given as

$$Y = AX + E, \quad (5.3)$$

with random noise  $E$ . In Bayesian inference,  $X$  is unknown, however information about  $X$  may be inferred using assumptions on  $X$  and  $E$ , and observations  $Y$ . More precisely, we consider a realization  $\mathbf{x}$  of  $X$  with prior probability density function (PDF)  $\pi_{\text{prior}} := \pi(\mathbf{x})$ , and a realization  $\mathbf{e}$  of  $E$  for the the observations according to the PDF  $\pi_{\text{noise}} := \pi(\mathbf{y}|\mathbf{x})$ .

The objective of Bayesian inference is to determine the PDF that describes the input parameters  $\mathbf{x}$  given noisy observations  $\mathbf{y}$ . Bayes' rule provides this relationship, i.e., the probability of  $\mathbf{x}$  conditioned on observations  $\mathbf{y}$ ,  $\pi_{\text{post}} := \pi(\mathbf{x}|\mathbf{y})$ , as

$$\pi(\mathbf{x}|\mathbf{y}) = \frac{\pi(\mathbf{y}|\mathbf{x})\pi(\mathbf{x})}{\pi(\mathbf{y})}. \quad (5.4)$$

In Bayesian inference, information regarding  $\pi(\mathbf{y}|\mathbf{x})$  and  $\pi(\mathbf{x})$  is available from assumed statistics on the prior and noise. For scenarios of interest to this work, the forward model is linear, and both the

observational noise and prior distribution are Gaussian, and thus the posterior distribution will be Gaussian as well. Let  $\mathbf{A} \in \mathbb{R}^{m \times N}$ , with  $m \ll N$ , denote the discrete linear parameter-to-observable operator, such that

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}, \quad (5.5)$$

where  $\mathbf{x} \in \mathbb{R}^N$  and  $\mathbf{y}, \mathbf{e} \in \mathbb{R}^m$ . As the components of (5.4) are not all known, the relation

$$\pi(\mathbf{x}|\mathbf{y}) \propto \pi(\mathbf{y}|\mathbf{x})\pi(\mathbf{x}) \quad (5.6)$$

dictates that the posterior is proportional to the product of the likelihood and the prior. Assuming the prior distribution is  $\mathcal{N}(\mathbf{0}, \mathbf{\Gamma}_{\text{prior}})$  and the noise is described with distribution  $\mathcal{N}(\mathbf{0}, \mathbf{\Gamma}_{\text{noise}})$ , (5.6) can be written as

$$\pi(\mathbf{x}|\mathbf{y}) \propto \exp\left(-\frac{1}{2}\|\mathbf{\Gamma}_{\text{noise}}^{-1/2}(\mathbf{A}\mathbf{x} - \mathbf{y})\|^2 - \frac{1}{2}\|\mathbf{\Gamma}_{\text{prior}}^{-1/2}\mathbf{x}\|^2\right), \quad (5.7)$$

where  $\|\cdot\|$  indicates the spectral norm. As  $\mathbf{A}$  is a linear map, the posterior mean can be determined via maximum *a posteriori* (MAP) estimation. The posterior mean is given as

$$\bar{\mathbf{x}}_{\text{post}} = \arg \min_{\mathbf{x}} \left( \frac{1}{2}\|\mathbf{\Gamma}_{\text{noise}}^{-1/2}(\mathbf{A}\mathbf{x} - \mathbf{y})\|^2 + \frac{1}{2}\|\mathbf{\Gamma}_{\text{prior}}^{-1/2}\mathbf{x}\|^2 \right), \quad (5.8)$$

and posterior covariance  $\mathbf{\Gamma}_{\text{post}}$  at the MAP point, as defined in (5.1), is

$$\mathbf{\Gamma}_{\text{post}} = \left( \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A} + \mathbf{\Gamma}_{\text{prior}}^{-1} \right)^{-1}.$$

Since  $m \ll N$ ,  $\mathbf{A}\mathbf{x} = \mathbf{y}$  is ill-posed, and thus regularization methods, e.g., MAP estimation, are used to calculate the posterior mean. A naïve approach to calculating  $\bar{\mathbf{x}}_{\text{post}}$  requires forming the posterior covariance in (5.1), which entails forming and inverting the  $N \times N$  Hessian matrix. For large problems this is infeasible, and thus an important area of research.

**Remark 5.2.1.** *In best practices, the prior covariance is a discretization-invariant operator [90]. For instance, consider  $\mathbf{\Gamma}_{\text{prior}}^{-1/2} = \sqrt{\alpha}\mathbf{L}$ , where  $\mathbf{L}$  is a discretization-invariant operator such as the Laplace operator. The action of the regularization term in (5.8) updated as  $\|\sqrt{\alpha}\mathbf{L}\mathbf{x}\|$  converges with respect to refinement of the discretization scheme used for the forward solver. That is, for*

$\mathbf{x} \in \mathbb{R}^N$ , as  $N$  is increased (i.e. the degrees of freedom of the numerical solver are increased), we see convergence in solutions  $\bar{\mathbf{x}}_{\text{post}}$  in (5.8). A discretization invariant operator should, in general, be used for the regularization operator – in this setting the posterior covariance – as it will allow for convergence with respect to refinement of the discretization scheme used for the forward solver.

### 5.3 Low-Rank Approximation for Posterior Covariance Estimation

As discussed previously, the direct calculation of  $\mathbf{\Gamma}_{\text{post}}$  requires the inversion of  $\mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A} + \mathbf{\Gamma}_{\text{prior}}^{-1}$ , which scales as  $\mathcal{O}(N^3)$ . For large  $N$ , this is infeasible. To mitigate this cost the authors of [40, 18] formulate a dimension reduction technique that exploits the low-rank structure of the prior preconditioned data misfit of the Hessian, given as

$$\mathbf{H} := \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A} \mathbf{\Gamma}_{\text{prior}}^{1/2}. \quad (5.9)$$

First, they reformulate the problem in (5.1) as (5.2)

$$\mathbf{\Gamma}_{\text{post}} = \mathbf{\Gamma}_{\text{prior}}^{1/2} (\mathbf{H} + \mathbf{I})^{-1} \mathbf{\Gamma}_{\text{prior}}^{1/2}.$$

Second, they compute the  $r$  dominant eigenvalues and associated eigenvectors of  $\mathbf{H}$  using the Lanczos algorithm (see Appendix C.1.1 and Appendix C.1.2 for details). This results in the rank  $r$  estimate of  $\mathbf{H}$

$$\mathbf{H} \approx \mathbf{V}_r \mathbf{\Lambda}_r \mathbf{V}_r^T, \quad (5.10)$$

where orthonormal columns of  $\mathbf{V}_r \in \mathbb{R}^{N \times r}$  and diagonal components of  $\mathbf{\Lambda}_r \in \mathbb{R}^{r \times r}$  are the first  $r$  eigenvectors and eigenvalues (respectively) of  $\mathbf{H}$ . Here we highlight that the full eigenvalue decomposition is given as  $\mathbf{H} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T$ , where  $\mathbf{V} \in \mathbb{R}^{N \times N}$  and  $\mathbf{\Lambda} \in \mathbb{R}^{N \times N}$ . In the final step, they use the Sherman-Morrison-Woodbury formula to form the rank  $r$  approximation

$$\mathbf{\Gamma}_{\text{post}} = \mathbf{\Gamma}_{\text{prior}}^{1/2} (\mathbf{H} + \mathbf{I})^{-1} \mathbf{\Gamma}_{\text{prior}}^{1/2} \approx \mathbf{\Gamma}_{\text{prior}}^{1/2} (\mathbf{I} - \mathbf{V}_r \mathbf{D}_r \mathbf{V}_r^T) \mathbf{\Gamma}_{\text{prior}}^{1/2}, \quad (5.11)$$

where  $\mathbf{D}_r = \text{diag}(\lambda_1/(\lambda_1 + 1), \dots, \lambda_r/(\lambda_r + 1))$ .

As this methodology reduces the problem from an  $N \times N$  matrix inversion to a rank  $r$  eigenvalue decomposition, the computational cost of forming the decomposition of (5.11) relies on the

technique for eigenvalue solve. While there are multiple methods to do such, several works consider matrix-free methods such as Lanczos (or randomized Lanczos) algorithm to identify columns of  $\mathbf{V}_r$  and entries of  $\mathbf{\Lambda}_r$ , as these are ideal for large-scale problems [40, 18, 19]. Specifically, these methods do not require forming  $\mathbf{H}$ , rather perform only matrix-vector products with each component of  $\mathbf{H}$ . For more information regarding these algorithms, see Appendix C.1.

Forming  $\mathbf{V}_r$  and  $\mathbf{\Lambda}_r$  require at least  $r$  forward and  $r$  adjoint solves of the high-fidelity model. This occurs by performing matrix-vector products with  $\mathbf{H}$ . For a vector  $\mathbf{q}_0$ , there are four major matrix-vector products that occur during the Lanczos algorithm: (i)  $\mathbf{q}_1 = \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{q}_0$ , (ii)  $\mathbf{q}_2 = \mathbf{A} \mathbf{q}_1$ , (iii)  $\mathbf{q}_3 = \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{q}_2$  where the product with  $\mathbf{\Gamma}_{\text{noise}}^{-1}$  often has negligible cost, and (iv)  $\mathbf{q}_4 = \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{q}_3$ . Items (i) and (iii) require matrix-vector products with the prior covariance. Depending on the choice of  $\mathbf{\Gamma}_{\text{prior}}$  this may or may not provide significant computational cost. Typically, one wants a discretization invariant operator, and it is suggested to consider  $\mathbf{\Gamma}_{\text{prior}}^{1/2}$  to be the Laplace operator or be a forward solve of an elliptic PDE as in [18] (see Remark 5.2.1). Items (ii) and (iii) are matrix-vector products containing forward and adjoint solves of the high-fidelity model, respectively. These two components account for the greatest computational cost.

The parameter-to-observable map  $\mathbf{A}$  is defined to be  $m \times N$  matrix; however, it contains a product of multiple matrices. For example, we may define  $\mathbf{A} := \mathbf{P} \mathbf{G}$  where  $\mathbf{G}$  is an  $N \times N$  matrix mapping parameters  $\mathbf{x}$  to solutions  $\mathbf{u}$  of the PDE, and  $\mathbf{P}$  is an  $m \times N$  permutation matrix that extracts specific coordinate points from the solution vector  $\mathbf{u}$  to form observations  $\mathbf{y}$ . In scenarios of interest, matrix-vector products  $\mathbf{G} \mathbf{x}$  are expensive, being equivalent to a forward solve of the fine grid model. As we will see in the numerical examples of Section 5.5, the forward PDE solver may in fact contain multiple matrix-vector products when time stepping is involved.

Finally, we comment on the accuracy of this approximation. Recall the full eigenvalue decomposition of the Hessian misfit is  $\mathbf{H} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T$ . Then, using the Sherman-Morrison-Woodbury formula, it can be shown that

$$\|(\mathbf{H} + \mathbf{I})^{-1} - (\mathbf{I} - \mathbf{V}_r \mathbf{D}_r \mathbf{V}_r^T)\| = \frac{\lambda_{r+1}}{1 + \lambda_{r+1}}, \quad (5.12)$$

where  $\|\cdot\|$  is the spectral norm, and  $\lambda_{r+1}$  is the  $r+1$  largest eigenvalue of  $\mathbf{H}$  (see Theorem 5.4.2 for details). We note this is not the error of the posterior covariance, as the left and right products with the prior covariance are absent; however, this error is valuable, as it may be determined numerically with access to the  $r+1$  eigenvalue only. In the numerical results section (Section 5.5), we calculate errors in this form, as we may compare the numerical errors with the error of (5.12).

#### 5.4 Low-Rank, Bi-Fidelity Approximation for Posterior Covariance Estimation

In this section we introduce our bi-fidelity formulation of the posterior covariance. This method follows a similar algorithm to the Nyström approximation as in [42, 141]; however, uses a multi-fidelity framework. Specifically, the Nyström approximation allows us to identify the right singular vectors  $\mathbf{V}$  via a mapping of the left singular vectors and associated singular values. We consider a similar method, but with a focus on utilizing the mapping for the coarse grid solution. Before discussing the procedure, we introduce the coarse grid mapping.

Assume that we have access to an  $m \times n$  coarse grid version of the parameter-to-observable map, denoted  $\mathbf{B}$ , that maps a reduced dimension (coarse) version of the parameter space to the observables, with  $n < N$ . As  $\mathbf{B}$  contains a coarse grid solver, both the inputs  $\tilde{\mathbf{x}}$  and outputs  $\tilde{\mathbf{y}}$  are considered to be low-fidelity approximations to the high-fidelity versions. More precisely,  $\tilde{\mathbf{x}} \in \mathbb{R}^n$  is the input parameter defined on a lower resolution space, and  $\tilde{\mathbf{y}} \in \mathbb{R}^m$  is the output of observations determined from the low-fidelity model. Define  $\tilde{\mathbf{\Gamma}}_{\text{prior}}^{1/2}$  to be a coarse version of  $\mathbf{\Gamma}_{\text{prior}}^{1/2}$ ; specifically,  $\tilde{\mathbf{\Gamma}}_{\text{prior}}$  is the prior covariance for the unknown parameters on the reduced dimensional space. We let  $\mathbf{\Gamma}_{\text{prior}}^{1/2}$  be a discretization-invariant operator, as discussed in Remark 5.2.1. We could think of a coarse version of the data misfit of the Hessian in (5.9) to be

$$\tilde{\mathbf{H}} = \tilde{\mathbf{\Gamma}}_{\text{prior}}^{1/2} \mathbf{B}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{B} \tilde{\mathbf{\Gamma}}_{\text{prior}}^{1/2};$$

however, in this case,  $\tilde{\mathbf{H}}$  would provide only information related to  $n$  unknowns of the coarse space. Instead, we consider using the left singular vectors and singular values of  $\mathbf{\Gamma}_{\text{noise}}^{-1/2} \mathbf{B} \tilde{\mathbf{\Gamma}}_{\text{prior}}^{1/2}$  for a bi-



fidelity approach to estimating the rank  $r$  eigenvalue decomposition of  $\mathbf{H}$ . The algorithm follows in three main steps: (i) identification of left singular vectors and singular values of the coarse grid version of mapping  $\mathbf{\Gamma}_{\text{noise}}^{-1/2} \mathbf{B} \tilde{\mathbf{\Gamma}}_{\text{prior}}^{1/2}$ , and (ii) lifting, and (iii) formation of posterior covariance via Sherman-Morrison-Woodbury formula.

In the first step, the left singular vectors and singular values of the coarse grid version of mapping are identified. This can be done in two ways. If  $n$  is small enough, a truncated SVD or randomized SVD [57, 84] can be done on the forward map such that

$$\mathbf{\Gamma}_{\text{noise}}^{-1/2} \mathbf{B} \tilde{\mathbf{\Gamma}}_{\text{prior}}^{1/2} \approx \tilde{\mathbf{U}}_r \tilde{\mathbf{\Sigma}}_r \tilde{\mathbf{V}}_r^T, \quad (5.13)$$

where  $\tilde{\mathbf{U}}_r$  is the  $m \times r$  matrix of left singular vectors,  $\tilde{\mathbf{V}}_r$  is the  $n \times r$  matrix of right singular vectors, and  $\tilde{\mathbf{\Sigma}}_r$  is the  $r \times r$  diagonal matrix containing largest  $r$  the singular values. Alternatively, a matrix-free approach such as Lanczos algorithm may be done on the product of the matrix in (5.13) with its transpose, i.e.,

$$\mathbf{\Gamma}_{\text{noise}}^{-1/2} \mathbf{B} \tilde{\mathbf{\Gamma}}_{\text{prior}} \mathbf{B}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} \approx \tilde{\mathbf{U}}_r \tilde{\mathbf{\Lambda}}_r \tilde{\mathbf{U}}_r^T, \quad (5.14)$$

where  $\tilde{\mathbf{\Lambda}}_r = \tilde{\mathbf{\Sigma}}_r \tilde{\mathbf{\Sigma}}_r^T$  is the  $r \times r$  matrix containing the coarse grid eigenvalues. Typically we are unable to perform direct SVD on the high-fidelity version of this mapping, thus a matrix-free approach is ideal for the low-fidelity mapping. Once this is complete, we have the truncated rank  $r$  left singular vectors  $\tilde{\mathbf{U}}_r$  ( $m \times r$ ) and eigenvalues  $\tilde{\mathbf{\Lambda}}_r$  ( $r \times r$ ).

In the second step,  $\tilde{\mathbf{U}}_r$  and  $\tilde{\mathbf{\Lambda}}_r$  are applied to the fine grid forward map to approximate the fine grid eigenvectors. Define

$$\hat{\mathbf{V}}_r := \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} \tilde{\mathbf{U}}_r (\tilde{\mathbf{\Lambda}}_r^{1/2})^+ \quad (5.15)$$

where  $^+$  denotes the pseudoinverse. From this definition, we can approximate the data misfit of the Hessian as

$$\hat{\mathbf{H}}_r := \hat{\mathbf{V}}_r \tilde{\mathbf{\Lambda}}_r \hat{\mathbf{V}}_r. \quad (5.16)$$

In the final step, Sherman-Morrison-Woodbury formula is applied when replacing the high-fidelity

Hessian misfit in (5.9) with the bi-fidelity version of (5.16) :

$$\mathbf{\Gamma}_{\text{prior}}^{1/2} \left( \hat{\mathbf{H}}_r + \mathbf{I} \right)^{-1} \mathbf{\Gamma}_{\text{prior}}^{1/2} = \mathbf{\Gamma}_{\text{prior}}^{1/2} \left( \mathbf{I} - \hat{\mathbf{V}}_r \hat{\mathbf{D}}_r \hat{\mathbf{V}}_r^T \right) \mathbf{\Gamma}_{\text{prior}}^{1/2}, \quad (5.17)$$

where  $\hat{\mathbf{D}}_r = (\mathbf{I} + \tilde{\mathbf{\Lambda}}_r \hat{\mathbf{V}}_r^T \hat{\mathbf{V}}_r)^{-1} \tilde{\mathbf{\Lambda}}_r$ .

**Remark 5.4.1.** *Given the formulation of  $\hat{\mathbf{V}}_r$  in (5.15), its columns are not orthogonal, resulting in  $\hat{\mathbf{V}}_r^T \hat{\mathbf{V}}_r \neq \mathbf{I}$  in the definition of  $\hat{\mathbf{D}}_r$ .*

In Section 5.3 the primary cost was composed of four matrix-vector products accounting for the majority of the computational cost. For this bi-fidelity procedure, we are improving the cost by performing items (i) and (ii) with coarse grid versions of the forward solvers (instead of fine grid solvers), and only items (iii) and (iv) require fine grid versions of the adjoint solvers. Depending on the cost difference in the coarse and fine grid solvers, as well as implementation for calculating the eigenvalue decompositions, we will see a range of possible cost reductions; regardless, the cost associated with the number of fine grid solves will be reduced by a factor of 2. Further discussion of cost comparisons is presented in Section 5.4.2.

As we will see in Section 5.5, this implementation is accurate for small values of the rank  $r$ . The success of this method relies on the fact that both fine grid and coarse grid parameter-to-observables maps extract values at the same coordinate points. The decompositions of these mappings allow us to identify the most important modes that describe the observations. While the coarse grid map may not accurately return observations, as long as it is able to identify similar modes to those of the fine grid mapping, we will see that this bi-fidelity approximation performs well.

#### 5.4.1 Theoretical Error Bound For Bi-Fidelity Approximation

For more analytical groundwork regarding the accuracy of this approximation, we discuss the theoretical error of the bi-fidelity approximation to the posterior covariance. This error bound relies on knowledge of the mapping between the fine grid and coarse grid version of the mapping,

specifically, that we can form a rank  $r$  mapping  $\mathbf{X}_r$ , where

$$\mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} = \mathbf{X}_r (\tilde{\mathbf{\Gamma}}_{\text{prior}}^{1/2} \mathbf{B}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2}) + \mathbf{E}, \quad (5.18)$$

and that  $\|\mathbf{E}\|$  small. Define the rank  $r$  mapping to between the low-fidelity and high-fidelity adjoint operators to be

$$\mathbf{X}_r = \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} (\tilde{\mathbf{V}}_r \tilde{\mathbf{\Sigma}}_r \tilde{\mathbf{U}}_r^T)^+, \quad (5.19)$$

where the  $\tilde{\mathbf{U}}_r \tilde{\mathbf{\Sigma}}_r \tilde{\mathbf{V}}_r^T$  forms the SVD of the coarse grid map as in (5.13). This construction of  $\mathbf{X}_r$  is specific to the low-fidelity and high-fidelity operators, and is exploited in the theoretical results to follow. We note that all theoretical error estimates require  $\tilde{\mathbf{\Sigma}}_r^+ \tilde{\mathbf{\Sigma}}_r = \mathbf{I}$ , and thus are applicable for rank  $r \leq \text{rank}(\tilde{\mathbf{\Sigma}})$ .

**Lemma 5.4.1.** *For rank  $r \leq \min(m, \text{rank}(\tilde{\mathbf{\Sigma}}))$ , the error of the mapping in (5.18) is*

$$\mathbf{E} = \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} (\mathbf{I} - \tilde{\mathbf{U}}_r \tilde{\mathbf{U}}_r^T). \quad (5.20)$$

*Proof.* From the definitions in (5.18) and (5.19), it follows that

$$\begin{aligned} \mathbf{E} &= \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} - \mathbf{X}_r (\tilde{\mathbf{\Gamma}}_{\text{prior}}^{1/2} \mathbf{B}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2}) \\ &= \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} - \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} (\tilde{\mathbf{V}}_r \tilde{\mathbf{\Sigma}}_r \tilde{\mathbf{U}}_r^T)^+ \tilde{\mathbf{\Gamma}}_{\text{prior}}^{1/2} \mathbf{B}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} \\ &= \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} - \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} (\tilde{\mathbf{U}}_r \tilde{\mathbf{\Sigma}}_r^+ \tilde{\mathbf{V}}_r^T) \tilde{\mathbf{V}}_r \tilde{\mathbf{\Sigma}}_r \tilde{\mathbf{U}}_r^T \\ &= \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} - \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} \tilde{\mathbf{U}}_r \tilde{\mathbf{U}}_r^T. \end{aligned}$$

□

**Lemma 5.4.2.** *Define  $\mathbf{H}$  as in (5.9),  $\hat{\mathbf{H}}_r$  as in (5.16), and  $\mathbf{E}$  as in (5.20). For  $r \leq \min(m, \text{rank}(\tilde{\mathbf{\Sigma}}))$  it follows that*

$$\mathbf{H} - \hat{\mathbf{H}}_r = \mathbf{E} \mathbf{\Gamma}_{\text{noise}}^{-1/2} \mathbf{A} \mathbf{\Gamma}_{\text{prior}}^{1/2}.$$

*Proof.* First  $\hat{\mathbf{H}}_r$  may be rewritten as,

$$\begin{aligned}\hat{\mathbf{H}}_r &= \hat{\mathbf{V}} \tilde{\mathbf{\Lambda}}_r \hat{\mathbf{V}}^T \\ &= \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} \tilde{\mathbf{U}}_r (\tilde{\mathbf{\Lambda}}_r^+)^{1/2} \tilde{\mathbf{\Lambda}}_r (\tilde{\mathbf{\Lambda}}_r^+)^{1/2} \tilde{\mathbf{U}}_r^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} \mathbf{A} \mathbf{\Gamma}_{\text{prior}}^{1/2} \\ &= \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} \tilde{\mathbf{U}}_r \tilde{\mathbf{U}}_r^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} \mathbf{A} \mathbf{\Gamma}_{\text{prior}}^{1/2}.\end{aligned}$$

Then it follows that the error is

$$\begin{aligned}\mathbf{H} - \hat{\mathbf{H}}_r &= \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1} \mathbf{A} \mathbf{\Gamma}_{\text{prior}}^{1/2} - \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} \tilde{\mathbf{U}}_r \tilde{\mathbf{U}}_r^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} \mathbf{A} \mathbf{\Gamma}_{\text{prior}}^{1/2} \\ &= \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} (\mathbf{I} - \tilde{\mathbf{U}}_r \tilde{\mathbf{U}}_r^T) \mathbf{\Gamma}_{\text{noise}}^{-1/2} \mathbf{A} \mathbf{\Gamma}_{\text{prior}}^{1/2} \\ &= \mathbf{E} \mathbf{\Gamma}_{\text{noise}}^{-1/2} \mathbf{A} \mathbf{\Gamma}_{\text{prior}}^{1/2}.\end{aligned}$$

□

**Lemma 5.4.3.** Define  $\mathbf{E}$  as in (5.20). For  $r \leq \min(m, \text{rank}(\tilde{\mathbf{\Sigma}}))$  it follows that

$$\mathbf{E} \mathbf{\Gamma}_{\text{noise}}^{-1/2} \mathbf{A} \mathbf{\Gamma}_{\text{prior}}^{1/2} = \mathbf{E} \mathbf{E}^T.$$

*Proof.* For  $r < m$ ,  $\mathbf{I} - \tilde{\mathbf{U}}_r \tilde{\mathbf{U}}_r^T$  is a projection, and it follows that

$$\begin{aligned}\mathbf{E} \mathbf{\Gamma}_{\text{noise}}^{-1/2} \mathbf{A} \mathbf{\Gamma}_{\text{prior}}^{1/2} &= \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} (\mathbf{I} - \tilde{\mathbf{U}}_r \tilde{\mathbf{U}}_r^T) \mathbf{\Gamma}_{\text{noise}}^{-1/2} \mathbf{A} \mathbf{\Gamma}_{\text{prior}}^{1/2} \\ &= \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} (\mathbf{I} - \tilde{\mathbf{U}}_r \tilde{\mathbf{U}}_r^T) (\mathbf{I} - \tilde{\mathbf{U}}_r \tilde{\mathbf{U}}_r^T) \mathbf{\Gamma}_{\text{noise}}^{-1/2} \mathbf{A} \mathbf{\Gamma}_{\text{prior}}^{1/2} \\ &= \mathbf{E} \mathbf{E}^T\end{aligned}$$

For  $r = m = \text{rank}(\tilde{\mathbf{\Sigma}})$ , it follows that  $\mathbf{E} = \mathbf{0}$ , and thus the equality holds. □

Finally, we present the main theorem of this work, which allows us to gauge the accuracy of the approximation. While this error is not comparable to that of the posterior covariance, it is of interest, as it is comparable to the theoretical error  $\lambda_{r+1}/(1 + \lambda_{r+1})$  as in (5.12).

**Theorem 5.4.1.** Define  $\mathbf{H}$  as in (5.9) and  $\hat{\mathbf{H}}_r$  as in (5.16). For  $r < m$  and  $r \leq \text{rank}(\tilde{\mathbf{\Sigma}})$ , the error of the rank  $r$  bi-fidelity approximation to  $(\mathbf{I} + \mathbf{H})^{-1}$  is bounded as

$$\|(\mathbf{I} + \mathbf{H})^{-1} - (\mathbf{I} + \hat{\mathbf{H}}_r)^{-1}\| \leq \|\mathbf{E}\|^2. \quad (5.21)$$

When  $r = m = \text{rank}(\tilde{\mathbf{\Sigma}})$  the error of the rank  $r$  approximation is

$$\|(\mathbf{I} + \mathbf{H})^{-1} - (\mathbf{I} + \hat{\mathbf{H}}_r)^{-1}\| = 0. \quad (5.22)$$

*Proof.* The error may be bounded as

$$\begin{aligned} \|(\mathbf{I} + \mathbf{H})^{-1} - (\mathbf{I} + \hat{\mathbf{H}}_r)^{-1}\| &= \|(\mathbf{I} + \mathbf{H})^{-1}((\mathbf{I} + \hat{\mathbf{H}}_r) - (\mathbf{I} + \mathbf{H}))(\mathbf{I} + \hat{\mathbf{H}}_r)^{-1}\| \\ &= \|(\mathbf{I} + \mathbf{H})^{-1}(\hat{\mathbf{H}}_r - \mathbf{H})(\mathbf{I} + \hat{\mathbf{H}}_r)^{-1}\| \\ &\leq \|(\mathbf{I} + \mathbf{H})^{-1}\| \|\mathbf{E}\mathbf{\Gamma}_{\text{noise}}^{-1/2}\mathbf{A}\mathbf{\Gamma}_{\text{prior}}^{1/2}\| \|(\mathbf{I} + \hat{\mathbf{H}}_r)^{-1}\| \\ &\leq \|\mathbf{E}\|^2, \end{aligned}$$

where the last two lines follow from Lemma 5.4.2 and Lemma 5.4.3; in addition, since  $\mathbf{H}$  and  $\hat{\mathbf{H}}_r$  are positive semi-definite, it follows that  $\|(\mathbf{I} + \mathbf{H})^{-1}\| \leq 1$  and  $\|(\mathbf{I} + \hat{\mathbf{H}}_r)^{-1}\| \leq 1$ .

For  $r = m = \text{rank}(\tilde{\mathbf{\Sigma}})$ , it follows that

$$\begin{aligned} \|(\mathbf{I} + \mathbf{H})^{-1} - (\mathbf{I} + \hat{\mathbf{H}}_r)^{-1}\| &\leq \|\mathbf{E}\mathbf{\Gamma}_{\text{noise}}^{-1/2}\mathbf{A}\mathbf{\Gamma}_{\text{prior}}^{1/2}\| \\ &= \|\mathbf{\Gamma}_{\text{prior}}^{1/2}\mathbf{A}^T\mathbf{\Gamma}_{\text{noise}}^{-1/2}(\mathbf{I} - \tilde{\mathbf{U}}_r\tilde{\mathbf{U}}_r^T)\mathbf{\Gamma}_{\text{noise}}^{-1/2}\mathbf{A}\mathbf{\Gamma}_{\text{prior}}^{1/2}\| \\ &= 0. \end{aligned}$$

Here we note that for  $r = m$ ,  $\tilde{\mathbf{U}}_r\tilde{\mathbf{U}}_r^T = \mathbf{I}$ , and the error becomes zero.  $\square$

**Remark 5.4.2.** When numerically testing the error bound of Theorem 5.4.1, error introduced according to machine precision must be taken into account. Specifically, the error for rank  $r = m = \text{rank}(\tilde{\mathbf{\Sigma}})$  will be approximately  $\lambda_1\epsilon_{\text{mach}}$ , where  $\lambda_1$  is the largest eigenvalue of  $\mathbf{H}$  and  $\epsilon_{\text{mach}}$  is machine epsilon.

The results of Theorem 5.4.1 show that if the error between the low-fidelity mapping and high-fidelity mapping is small, then the error of the rank  $r$  bi-fidelity approximation to  $(\mathbf{I} + \mathbf{H})^{-1}$  will be small; however, this is not a necessary condition for an accurate approximation.

A final theorem is provided to indicate when this bi-fidelity approximation will have the same error as the high-fidelity low-rank approximation. Recall the rank  $r$  truncated SVD of the coarse

grid mapping in (5.13), given as  $\mathbf{\Gamma}_{\text{noise}}^{-1/2} \mathbf{B} \tilde{\mathbf{\Gamma}}_{\text{prior}}^{1/2} \approx \tilde{\mathbf{U}}_r \tilde{\mathbf{\Sigma}}_r \tilde{\mathbf{V}}_r^T$ . We define the SVD for the fine grid mapping to be

$$\mathbf{\Gamma}_{\text{noise}}^{-1/2} \mathbf{A} \mathbf{\Gamma}_{\text{prior}}^{1/2} := \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T, \quad (5.23)$$

where  $\mathbf{U}$  and  $\mathbf{V}$  are orthonormal matrices of sizes  $m \times m$  and  $N \times N$ , respectively, and  $\mathbf{\Sigma}$  is the  $m \times N$  matrix containing the singular values. Note, this agrees with the definition in (5.10) with  $\mathbf{\Lambda} = \mathbf{\Sigma}^T \mathbf{\Sigma}$ , the matrix of eigenvalues for the Hessian misfit.

**Theorem 5.4.2.** *Given the definition of  $\mathbf{H}$  and  $\hat{\mathbf{H}}_r$  as well as the high-fidelity and low-fidelity map SVDs provided in (5.23) and (5.13), suppose that  $\mathbf{U}^T \tilde{\mathbf{U}}_r = [\mathbf{I} \mid \mathbf{0}]^T$ . Then the error of the rank  $r \leq \min(m, \text{rank}(\tilde{\mathbf{\Sigma}}))$  bi-fidelity approximation agrees with the standard low-rank error as in (5.12), namely,*

$$\|(\mathbf{I} + \mathbf{H})^{-1} - (\mathbf{I} + \hat{\mathbf{H}}_r)^{-1}\| = \frac{\lambda_{r+1}}{1 + \lambda_{r+1}}, \quad (5.24)$$

where  $\lambda_{r+1}$  is the  $r + 1$  largest eigenvalue of  $\mathbf{H}$ .

*Proof.* First we consider the SVD of  $\mathbf{H}$  and  $\hat{\mathbf{H}}_r$ . Recall the SVD of  $\mathbf{H}$  is provided in (5.10), and from (5.13)-(5.15), it can be shown that

$$\begin{aligned} \hat{\mathbf{H}}_r &= \mathbf{\Gamma}_{\text{prior}}^{1/2} \mathbf{A}^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} \tilde{\mathbf{U}}_r (\tilde{\mathbf{\Lambda}}_r^+)^{1/2} \tilde{\mathbf{\Lambda}}_r (\tilde{\mathbf{\Lambda}}_r^+)^{1/2} \tilde{\mathbf{U}}_r^T \mathbf{\Gamma}_{\text{noise}}^{-1/2} \mathbf{A} \mathbf{\Gamma}_{\text{prior}}^{1/2} \\ &= \mathbf{V} \mathbf{\Sigma} \mathbf{U}^T \tilde{\mathbf{U}}_r \tilde{\mathbf{U}}_r^T \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \\ &= \mathbf{V}_r \mathbf{\Lambda}_r \mathbf{V}_r^T. \end{aligned}$$

Then the bi-fidelity approximation to the data misfit of the Hessian  $\mathbf{H}$  is equivalent to the rank  $r$  approximation. Using the Sherman-Morrison-Woodbury formula, it follows that

$$(\mathbf{I} + \mathbf{H})^{-1} = \mathbf{I} - \mathbf{V}(\mathbf{I} + \mathbf{\Lambda})^{-1} \mathbf{\Lambda} \mathbf{V}^T.$$

Then the error may be expressed as

$$\begin{aligned}
\|(\mathbf{I} + \mathbf{H})^{-1} - (\mathbf{I} + \hat{\mathbf{H}}_r)^{-1}\| &= \|\mathbf{I} - \mathbf{V}(\mathbf{I} + \mathbf{\Lambda})^{-1}\mathbf{\Lambda}\mathbf{V}^T - (\mathbf{I} - \mathbf{V}_r(\mathbf{I} + \mathbf{\Lambda}_r)^{-1}\mathbf{\Lambda}_r\mathbf{V}_r^T)\| \\
&= \|\mathbf{V}_r(\mathbf{I} + \mathbf{\Lambda}_r)^{-1}\mathbf{\Lambda}_r\mathbf{V}_r^T - \mathbf{V}(\mathbf{I} + \mathbf{\Lambda})^{-1}\mathbf{\Lambda}\mathbf{V}^T\| \\
&= \|\mathbf{V}^T\mathbf{V}_r(\mathbf{I} + \mathbf{\Lambda}_r)^{-1}\mathbf{\Lambda}_r\mathbf{V}_r^T\mathbf{V} - (\mathbf{I} + \mathbf{\Lambda})^{-1}\mathbf{\Lambda}\| \\
&= \|[\mathbf{I} \mid \mathbf{0}]^T(\mathbf{I} + \mathbf{\Lambda}_r)^{-1}\mathbf{\Lambda}_r[\mathbf{I} \mid \mathbf{0}] - (\mathbf{I} + \mathbf{\Lambda})^{-1}\mathbf{\Lambda}\| \\
&= \frac{\lambda_{r+1}}{1 + \lambda_{r+1}}.
\end{aligned}$$

□

#### 5.4.2 Comparison of Computational Cost

In this section we discuss the cost reduction when using this bi-fidelity approximation for the posterior covariance. Overall, the primary cost reduction of using this method, in comparison to the standard high-fidelity, low-rank method, is that it requires half as many solves with the high-fidelity model. Thus the total cost reduction is based on the difference in cost of employing the low-fidelity and high-fidelity mapping, as well as the technique with which the eigenvalue decomposition is calculated.

Table 5.1 provides an approximation of the number of floating-point operations (FLOPs) for four methods: randomized Lanczos, randomized SVD, bi-fidelity with randomized Lanczos, and bi-fidelity with randomized SVD. For further details regarding these approximations, i.e., FLOP count totals, see Appendix C.2. For a low-fidelity mapping with  $n$  deterministic degrees of freedom, we denote the cost of a single forward solve as  $\mathcal{C}_F(n)$  and the cost of a single adjoint solve as  $\mathcal{C}_A(n)$ . For a high-fidelity mapping with  $N$  deterministic degrees of freedom, we denote the cost of a single forward solve as  $\mathcal{C}_F(N)$  and the cost of a single adjoint solve as  $\mathcal{C}_A(N)$ . From Table 5.1 we see that randomized Lanczos algorithm is more computationally expensive than randomized SVD; in particular, both require  $r$  forward and  $r$  adjoint solves. However, randomized Lanczos requires about  $8r^2N$  flops to perform steps related to randomization, while randomized SVD requires  $2r^2N$  steps related to randomization. The bi-fidelity approaches perform the eigenvalue decomposition

Method	Approximate FLOPs
Randomized Lanczos	$\approx 8r^2N + r\mathcal{C}_F(N) + r\mathcal{C}_A(N)$
Randomized SVD	$\approx 2r^2N + r\mathcal{C}_F(N) + r\mathcal{C}_A(N)$
Bi-Fidelity with Randomized Lanczos	$\approx r^2N + rN + r\mathcal{C}_A(N) + 8r^2n + r\mathcal{C}_F(n) + r\mathcal{C}_A(n)$
Bi-Fidelity with Randomized SVD	$\approx r^2N + rN + r\mathcal{C}_A(N) + 2r^2n + r\mathcal{C}_F(n) + r\mathcal{C}_A(n)$

Table 5.1: Approximate Number of FLOPs for Different Low-Rank Approaches

on the low-fidelity mappings, and thus costs related to the eigenvalue decompositions scale with the cost of a coarse grid solve as opposed to the cost of a fine grid solve. Beyond the eigenvalue decompositions, both bi-fidelity approaches require  $r$  adjoint solves with the high-fidelity mapping (but not any forward solves with the high-fidelity mapping). Additional costs of order  $N$  relate to the cost to form  $\hat{\mathbf{V}}_r$  and  $\hat{\mathbf{D}}_r$  for the posterior covariance estimates.

While the cost differences between the high-fidelity and bi-fidelity methods rely on the difference between  $n$  and  $N$ , it is clear that the cost difference between randomized Lanczos and bi-fidelity with randomized Lanczos will be larger than the cost difference between randomized SVD and bi-fidelity with randomized SVD. This is due to randomized Lanczos requiring more FLOPs than randomized SVD. However, we note, that the randomized SVD method here does not incorporate power iteration, which may improve the accuracy [113, 112, 57]. Additionally, we don't include oversampling in these randomized methods [84, 57]. Including either power iteration or oversampling would result in further cost reduction when using our bi-fidelity approximation. Cost comparisons for specific examples will be presented in the Section 5.5.



## 5.5 Numerical Performance of Bi-Fidelity Approximation to Posterior Covariance

In this section we show the promise of this low-rank, bi-fidelity approximation by comparing its performance with that of the high-fidelity and low-fidelity versions. The aim of these results is to show that the accuracy of the approximation is similar to that of the high-fidelity formulation for smaller ranks, while maintaining a reduced cost.

For the primary example, we consider a two-dimensional advection-diffusion problem with known variable diffusion and unknown initial temperature  $u_0$ , given by the PDE

$$\begin{cases} \frac{\partial u}{\partial t} - \nabla \cdot (a \nabla u) + \mathbf{v} \cdot \nabla u = 0 & \text{in } D \times [0, T] \\ u = u_0 & \text{in } D \times \{t = 0\}, \\ u = 0 & \text{in } \partial D \times [0, T], \\ \frac{\partial u}{\partial n} = g & \text{in } \partial D \times [0, T], \end{cases} \quad (5.25)$$

where  $D = [0, 1]^2$  is the spatial domain,  $u$  is the temperature,  $\mathbf{v} = (1, 1)$  is the velocity of the advection term,  $g$  enforces continuity at the boundary, and  $T = 0.07$  is the final time. The diffusion  $a$  is a realization of a log-normal random field given as

$$a(\mathbf{x}, \boldsymbol{\xi}) = \bar{a} + \exp(G(\mathbf{x}, \boldsymbol{\xi})), \quad \mathbf{x} \in D,$$

with  $G(\mathbf{x}, \boldsymbol{\xi})$  a Gaussian random field represented by the Karhunen-Loève expansion

$$G(\mathbf{x}, \boldsymbol{\xi}) = \sum_{i=1}^d \sqrt{\lambda_i} \phi_i(\mathbf{x}) \xi_i,$$

where the random variables  $\xi_i$  are independent standard Gaussian and  $\lambda_i$  are the  $d$  largest eigenvalues with corresponding eigenfunctions  $\phi_i(\mathbf{x})$  determined by the Gaussian covariance function

$$\mathcal{K}(\mathbf{x}_1, \mathbf{x}_2) = \sigma^2 \exp\left(-\frac{\|\mathbf{x}_1 - \mathbf{x}_2\|_2^2}{\ell}\right).$$

For this example, we set  $\bar{a} = 0$ ,  $d = 32$ ,  $\sigma^2 = 0.1$ , and  $\ell = 0.05$ . Figure 5.1 (a) displays the realization of this log-normal random field used to represent the variable diffusion.

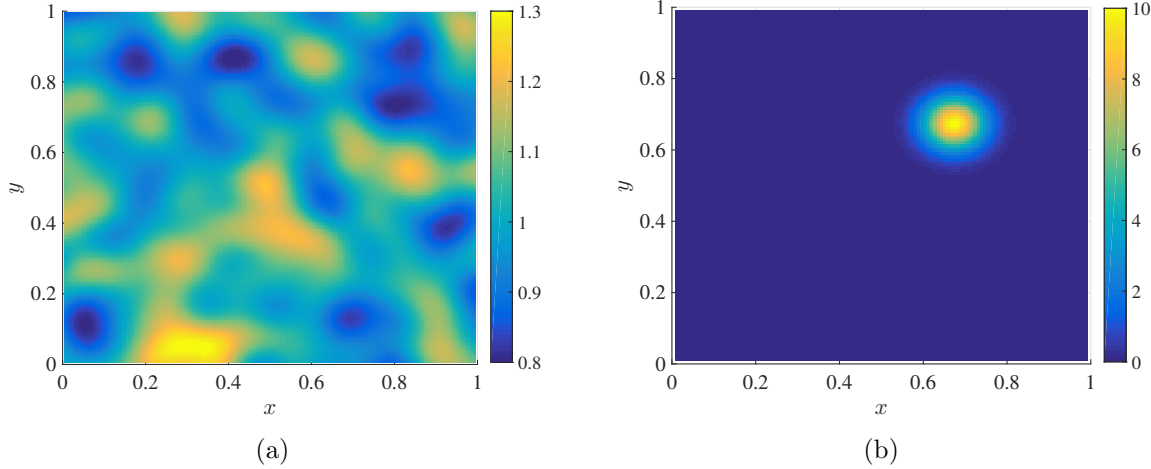


Figure 5.1: (a) Diffusion field over the physical domain  $D$ . (b) True initial value of  $\mathbf{u} = \mathbf{u}_0$  over  $D$  at time  $t = 0$ .

In this example, the forward parameter-to-observables map is the operator that maps the unknown initial condition  $u_0$  to the set of point observations of the solution  $u$ . Figure 5.1 (b) provides the image of  $\mathbf{u}_0$ , which is formed as a Gaussian peak at  $(0.67, 0.67)$ , with magnitude 10 and correlation length of  $\ell = 0.005$ .

In a numerical context, we consider finite-dimensional approximations. For the forward parameter-to-observable map, we consider the mapping  $\mathbf{A} = [\mathbf{P}\mathbf{G}_1 \mid \mathbf{P}\mathbf{G}_2 \mid \mathbf{P}\mathbf{G}_3]^T \mathbf{u}_0$ , where  $\mathbf{u}_0$  is a vector of the initial temperature over the spatial domain. In practice, we fix the unknown, and have access to observations  $\mathbf{y}$ , such that

$$\mathbf{y} = \mathbf{A}\mathbf{u}_0 + \mathbf{e}, \quad (5.26)$$

where the random variable  $\mathbf{e}$  indicates that the observations contain noise that deviates from the true solution. For the map  $\mathbf{A}$ , the forward solver  $\mathbf{G}_k \in \mathbb{R}^{N \times N}$  with  $k = 1, 2$ , or  $3$  performs finite differencing of explicit advection and implicit diffusion, such that, for a single time step, we have

$$\mathbf{u}^{(1)} = \left( \mathbf{I} - \frac{h_t}{h_x^2} \tilde{\mathbf{L}} \right)^{-1} \left( \mathbf{I} - \frac{h_t}{h_x} \mathbf{C} \right) \mathbf{u}_0$$

where  $h_x$  is the grid spacing in  $D$  (with uniform spacing),  $h_t$  is the time step size,  $\tilde{\mathbf{L}}$  is the Laplacian-like operator that incorporates the variable diffusion, and  $\mathbf{C}$  is the two-dimensional

upwind difference operator that provides the advection. The operator  $\tilde{\mathbf{L}}$  may be represented with the stencil

$$\frac{1}{h_x^2} \begin{bmatrix} & & a_{i,j-1/2} & & \\ a_{i-1/2,j} & -(a_{i,j-1/2} + a_{i-1/2,j} + a_{i+1/2,j} + a_{i,j+1/2}) & a_{i+1/2,j} & & \\ & & a_{i,j+1/2} & & \end{bmatrix}, \quad (5.27)$$

where  $a_{ij}$  corresponds to the value of the diffusion coefficient  $a(\mathbf{x}, \boldsymbol{\xi})$  at point  $\mathbf{x} = (ih_x, jh_x)$  in the spatial domain  $D$ , where  $i, j = 1, \dots, \sqrt{N}$ . Neumann boundary conditions are enforced by forming ghost nodes exterior to the boundary. These values are set to be the negative value of the node opposite the boundary (directly interior to the boundary). Then we define the forward PDE operator for this test case with  $n_k$  time steps to be

$$\mathbf{G}_k = \left[ \left( \mathbf{I} - \frac{h_t}{h_x^2} \tilde{\mathbf{L}} \right)^{-1} \left( \mathbf{I} - \frac{h_t}{h_x} \mathbf{C} \right) \right]^{n_k},$$

where  $0 < n_1 < n_2 < n_3 = 45$ . The solution to this forward solver at  $T = 0.07$  is provided in Figure 5.2. The observation matrix  $\mathbf{P}$  extracts  $m_0$  points in the spatial domain, and is applied at  $k = 3$  snapshots in time, making  $m = 3m_0$ . Note, we keep  $m$  in this form, as we will consider how the results change with different observation scenarios, specifically, increasing the value of  $m_0$ . The observational data  $\mathbf{y}$  is derived from this forward mapping with additive noise, as in (5.26).

**Remark 5.5.1.** *In practice the operators  $[\mathbf{P}\mathbf{G}_1 \mid \mathbf{P}\mathbf{G}_2 \mid \mathbf{P}\mathbf{G}_3]^T$  and  $\mathbf{G}_i$  are not directly formed.*

For the low-fidelity version of the parameter-to-observable map, we consider  $\mathbf{B}$ , as discussed in Section 5.4. This map, defined as  $\mathbf{B} = [\tilde{\mathbf{P}}\tilde{\mathbf{G}}_1 \mid \tilde{\mathbf{P}}\tilde{\mathbf{G}}_2 \mid \tilde{\mathbf{P}}\tilde{\mathbf{G}}_3]^T \in \mathbb{R}^{m \times n}$ , applies a coarse grid version of  $\mathbf{G}_i$ , denoted  $\tilde{\mathbf{G}}_i \in \mathbb{R}^{n \times n}$ , for  $i = 1, 2$ , and 3 and with  $n < N$  degrees of freedom in the spatial domain, to a coarse version of the initial condition, and then a corresponding permutation matrix  $\tilde{\mathbf{P}} \in \mathbb{R}^{m_0 \times n}$  extracts the solution at the same coordinate points as  $\mathbf{P}$ .

### 5.5.1 Test Case # 1: $m_0 = 9$

For the first test case, we consider the scenario where we have a  $3 \times 3$  grid of sensors uniformly distributed over the spatial domain ( $m_0 = 9$ ) and use three time snapshots of data, resulting in

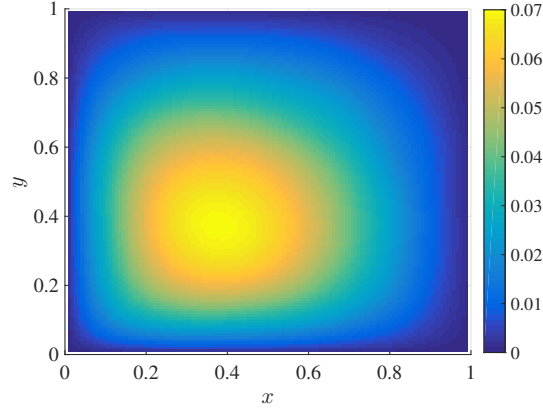


Figure 5.2: Solution  $\mathbf{u}$  at final time  $t = T$ , when using a reference mesh grid of size  $128 \times 128$ .

$m = 27$ . For the high-fidelity mapping, we consider a fine grid resolution with mesh size  $33 \times 33$ , and for the low-fidelity mapping, we consider a coarse grid resolution with mesh size  $9 \times 9$ , both of which are uniform. Thus  $N = 961$  and  $n = 49$  for solvers with zero boundary conditions. Figure 5.3 provides non-boundary data for (a) the fine grid and (b) the coarse grid meshes with red “x” markers indicating the sensor points, or more precisely, the points at which the data is extracted. The surface magnitudes display the final solution of the PDE in (5.25) for reference.

The noise of the observations,  $\mathbf{e}$ , is defined to be a realization of  $\mathcal{N}(\mathbf{0}, \mathbf{\Gamma}_{\text{noise}})$ , where the noise covariance is a diagonal matrix with entries  $\sigma_{\text{noise}}^2 \mathbf{A} \mathbf{u}_0$  with  $\sigma_{\text{noise}} = 0.03$ . This value of  $\sigma_{\text{noise}}$  assumes that the observational data has a 3% coefficient of variation. For the prior covariance, we define  $\mathbf{\Gamma}_{\text{prior}}^{-1} = \alpha \mathbf{L} \mathbf{L}^T$ , where  $\mathbf{L}$  is the Laplace operator. The regularization problem in (5.8) now contains the regularization term  $\alpha \|\mathbf{L} \mathbf{u}\|^2$ . Recall from Remark 5.2.1, that this is a discretization-invariant operator. Best practice for selecting  $\alpha$  is to inspect the L-curve when plotting the error term of (5.8) given as  $\|\mathbf{\Gamma}_{\text{noise}}^{-1/2}(\mathbf{A} \mathbf{u}_\alpha - \mathbf{y})\|$  against the magnitude of the regularized solution given as  $\|\mathbf{u}_\alpha\|$  [63, 64]. The curve provided in Figure 5.4 indicates that  $\alpha = 0.07$  is good choice for the regularization parameter.

For this example, we will consider two values of  $\alpha$  as the regularization parameter. The first,  $\alpha = 0.07$ , is determined via the L-curve in Figure 5.4, and the second,  $\alpha = 500$ , is to understand how increasing the regularization term affects the accuracy of the bi-fidelity approximation. Figure 5.5

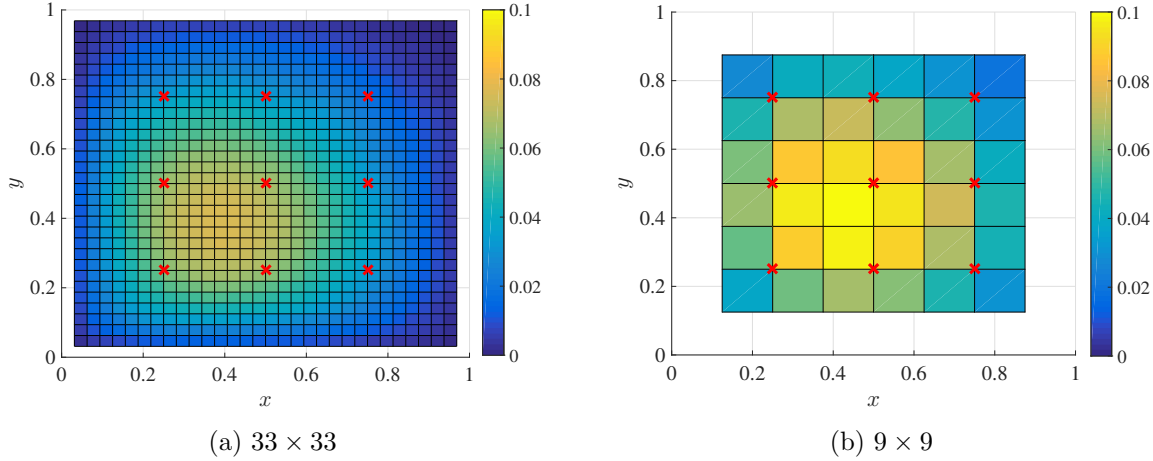


Figure 5.3: Numerical solutions (excluding boundary) to the PDE in (5.25) for (a) the fine grid mesh and (b) the coarse grid mesh. Red “x” markers indicate where observations are extracted.

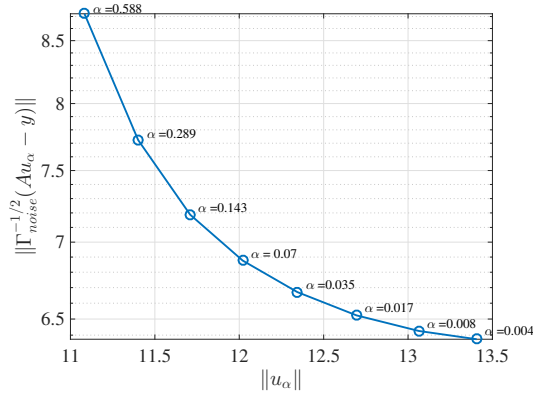


Figure 5.4: L-curve comparing the error term and regularization term of (5.8). Optimal point of  $\alpha = 0.07$  selected as balances the two terms.

provides the decay of the normalized eigenvalues for the high-fidelity and low-fidelity versions of the prior preconditioned data misfit of the Hessian for (a)  $\alpha = 0.07$  and (b)  $\alpha = 500$ . The rapid decay of eigenvalues of Figure 5.5 shows that both systems are low rank. Figure 5.5 also shows that the high-fidelity and low-fidelity mappings have a similar decay of eigenvalues for smaller ranks. From this, we see that the first 10 – 15 modes capture the most relevant information of the data, and thus we will likely only need rank  $r$  much less than  $m$  for the low-rank estimates.

Next we compare the relative error of the posterior covariance approximations determined by the low-rank, high-fidelity approach as in Section 5.3 and the low-rank, bi-fidelity approach as

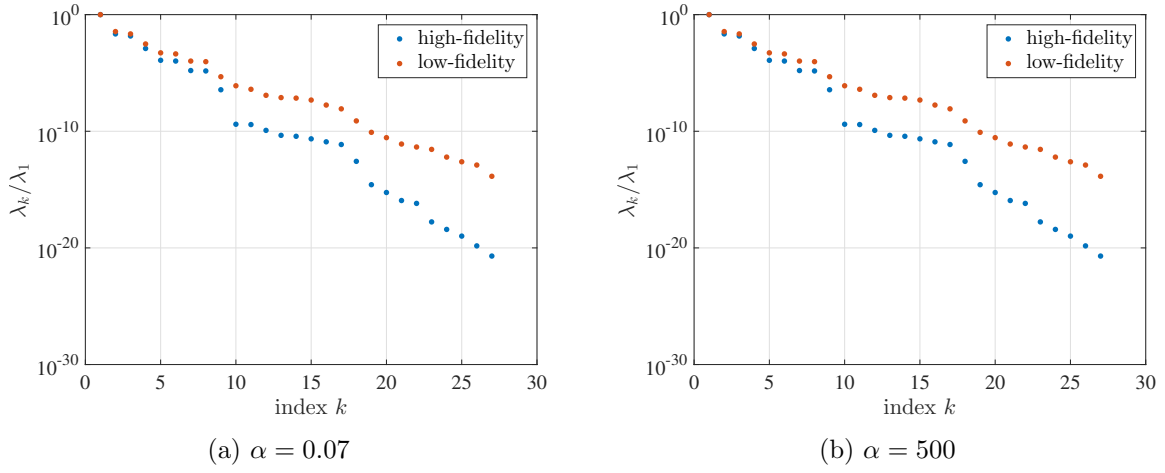


Figure 5.5: Decay of the eigenvalues of the data misfit of the Hessian for regularization parameter values (a)  $\alpha = 0.07$  and (b)  $\alpha = 500$ . High-fidelity refers the eigenvalues  $\mathbf{\Lambda}_r$ , as in (5.10), and low-fidelity refers to eigenvalues  $\tilde{\mathbf{\Lambda}}_r$ , as in (5.14).

in Section 5.4. The relative error is calculated as

$$\frac{\|\mathbf{\Gamma}_{\text{prior}}^{1/2}(\mathbf{I} + \mathbf{H})^{-1}\mathbf{\Gamma}_{\text{prior}}^{1/2} - \mathbf{\Gamma}_{\text{prior}}^{1/2}(\mathbf{I} - \mathbf{V}_r\mathbf{D}_r\mathbf{V}_r^T)\mathbf{\Gamma}_{\text{prior}}^{1/2}\|}{\|\mathbf{\Gamma}_{\text{prior}}^{1/2}(\mathbf{I} + \mathbf{H})^{-1}\mathbf{\Gamma}_{\text{prior}}^{1/2}\|}. \quad (5.28)$$

where  $\mathbf{\Gamma}_{\text{prior}}^{1/2} = \sqrt{\alpha}\mathbf{L}$ , and the matrices  $\mathbf{V}_r$  and  $\mathbf{D}_r$  are the rank  $r$  approximations determined from either Section 5.3 for the low-rank, high-fidelity approach, or Section 5.4 for the low-rank, bi-fidelity approach. For the bi-fidelity approximation error bound (see Theorem 5.4.1) we calculate the bound as

$$\frac{\|\mathbf{E}\|^2\|\mathbf{\Gamma}_{\text{prior}}^{1/2}\|^2}{\|\mathbf{\Gamma}_{\text{prior}}^{1/2}(\mathbf{I} + \mathbf{H})^{-1}\mathbf{\Gamma}_{\text{prior}}^{1/2}\|}, \quad (5.29)$$

where the error  $\mathbf{E}$  is defined in (5.20). As state-of-the-art is performing randomized methods, we look at two result types: direct SVD and randomized methods for calculating the eigenvalue decompositions. These results are provided in Figure 5.6, where the left images use direct SVD, and the right images use randomized Lanczos and randomized SVD for the low-rank and bi-fidelity approaches; these results all provide the decay of error as a function of rank  $r$ . Figure 5.6 (a) and (b) provide these results for the scenario of  $\alpha = 0.07$  (as determined from the L-curve), where the errors for the direct versus randomized methods behave quite similarly. When using direct SVD to calculate the eigenvalue decompositions, as in Figure 5.6 (a), we see an agreement between the low-rank (high-fidelity) method and the bi-fidelity approximation for ranks  $r < 8$ . At rank  $r = 8$  the low-rank approximation has a relative error of 0.002 and the bi-fidelity approximation has an error of 0.01. Thus if we require a 1% accuracy, we could apply either method. Figure 5.6 (b) shows the performance when using randomized methods. At rank  $r = 8$  we have 0.03 error for the randomized Lanczos, an error of 0.02 for both bi-fidelity approximations, and an error of 0.003 for randomized SVD. These errors are slightly larger than the direct SVD approaches; however, for larger problems (large  $N$ ), we would not be able to apply direct SVD, so it is necessary to observe how randomized methods behave. For both Figure 5.6 (a) and (b) the error bound from (5.29) is provided for comparison. For  $r < 27$  the bound is between  $\mathcal{O}(10^3)$  and  $\mathcal{O}(10^6)$  larger than the calculated error, showing that it does not provide a tight bound. At  $r = 27$  the bound behaves more like an error estimate, as the theoretical error is zero; this is because the numerical error relies

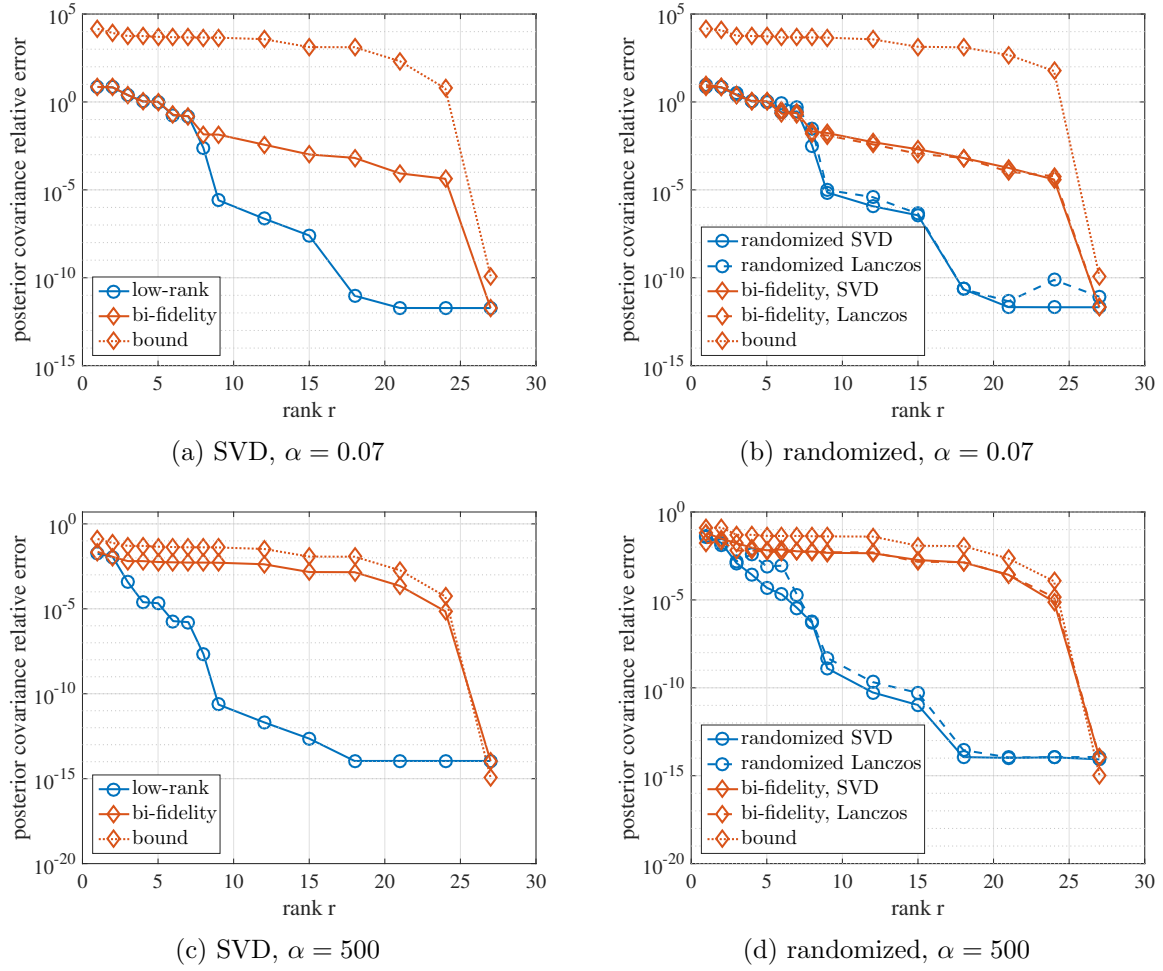


Figure 5.6: Relative error of the low-rank and low-rank, bi-fidelity approximations as in (5.28) for the scenarios (a)  $\alpha = 0.07$  and (b)  $\alpha = 500$ . Plots on the left correspond to the use of direct SVD to calculate the eigenvalue decompositions, and plots on the right correspond to using randomized approaches.

on machine precision (see Remark 5.4.2).

We consider how the errors and error bound behave for larger values of  $\alpha$ , specifically  $\alpha = 500$ , in Figure 5.6 (c) and (d). With an increase in  $\alpha$ , we observe an overall reduction in the error when using low-rank methods. Figure 5.6 (c) displays the error calculations when using direct SVD. At  $r = 3$ , the error of the low-rank approximation is 0.0003 and bi-fidelity approximations is 0.006, while the error bound 0.07. Thus we observe a significantly tighter error bound than for smaller values of  $\alpha$ . For  $r > 3$  the error of the low-rank approach continues to decay while the



error of the bi-fidelity approximation flattens. Thus for small  $r \leq 3$ , where we desire about three digits of accuracy, the bi-fidelity method provides a good approximation. Figure 5.6 (d) provides the errors when using randomized approaches. At rank  $r = 3$  randomized SVD and randomized Lanczos have an error of 0.001, bi-fidelity with randomized SVD has an error of 0.01, and bi-fidelity with randomized Lanczos an error of 0.006. When using randomized methods, the error bound is 0.07 at  $r = 3$ . Similar to the previous results, performing direct SVD results in a more accurate approximation, and the bi-fidelity approach performs similarly to low-rank approaches for small  $r$ . This indicates that for low-rank systems, we may obtain accurate enough approximations to estimate the posterior covariance with reduced cost. After determining how this approximation affects the estimation of the posterior mean, we will address the cost reduction of this method.

Next we consider the construction of the posterior mean by three different low-rank approaches; namely, the standard low-rank (high-fidelity) method, our bi-fidelity approach, and using a low-fidelity version of the low-rank method, for identifying the coordinates of the maximum point of  $u_0$ , located at  $(0.67, 0.67)$ . These errors are calculated using direct SVD. While this is not the primary task of this work, it is essential to address, as it shows that the high-fidelity model is more accurate than the low-fidelity model, and thus necessary for use, and that the accuracy of the bi-fidelity approximation is closer to that of the high-fidelity solution than the low-fidelity solution. Figure 5.7 displays the error of identifying this coordinate point, specifically, the Euclidean error between the true maximum point coordinate and the coordinate of the maximum point as determined from the posterior mean calculation in (5.8), as a function of the rank  $r$ .

In Figure 5.7 (a), which uses  $\alpha = 0.07$  for the prior covariance, we see that the bi-fidelity error more closely aligns with the low-rank (high-fidelity) method for rank  $r > 2$ . For rank  $r > 4$ , the bi-fidelity and low-rank (high-fidelity) errors are about 0.02, while the low-fidelity error is about 0.1. In Figure 5.7 (b), which uses  $\alpha = 500$  for the prior covariance, this error difference isn't as exaggerated. The bi-fidelity and low-rank (high-fidelity) errors are about 0.19, while the low-fidelity error is about 0.24. We make two conclusions from these results. First, we see from both sets of results that the low-rank (high-fidelity) method is more accurate than the low-fidelity, low-rank

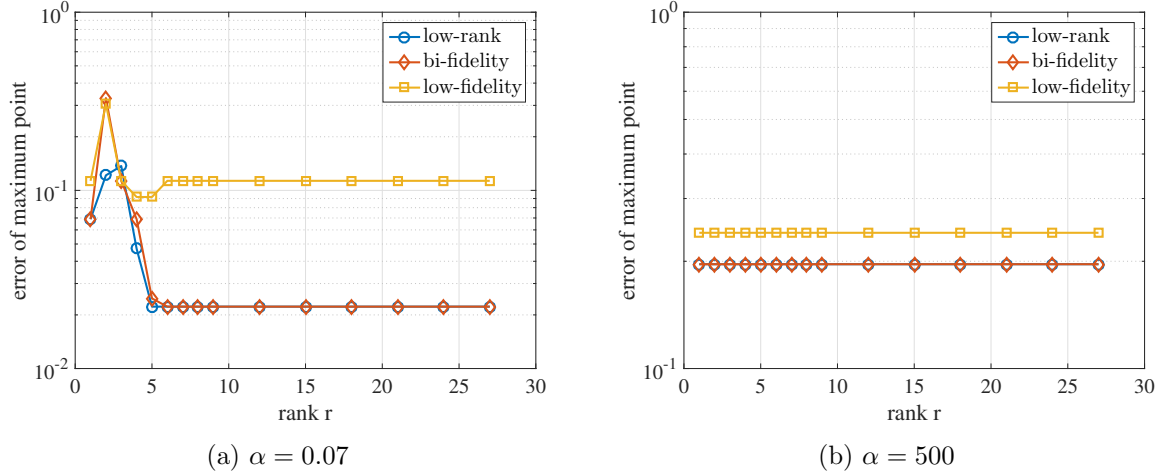


Figure 5.7: Euclidean distance between the true maximum point coordinate  $(0.67, 0.67)$  and the maximum point coordinate as calculated from the posterior mean (via MAP estimation) in (5.8). Three methods are compared: low-rank (high-fidelity), new bi-fidelity, and low-rank, low-fidelity approaches, as a function of rank  $r$ .

method, which utilizes the coarse grid solver; this shows that we will have a worse error with the coarse grid solver, and thus need to use the fine grid solver for improved accuracy. Second, the bi-fidelity approximation of the posterior covariance may be used in place of the low-rank (high-fidelity) approximation while maintaining the same order of accuracy for estimating the maximum point of  $u_0$ .

For the final result, we compare the number of FLOPs required for form the low-rank approximations. These estimates are based on treating the forward and adjoint solver with  $N$  degrees of freedom with cost  $\mathcal{C}_F(N) = \mathcal{C}_A(N) = 8Nn_t$ , where  $n_t$  is the number of time steps in the numerical solver. All numerical results are presented with  $n_t = 45$ ; however, for comparison, we present cost comparisons with  $n_t = 1$  as well to show the range of possible cost improvements. Figure 5.8 provides the FLOP calculation as discussed in Section 5.4.2 (and Appendix C.2). Figure 5.8 (a) provides the comparison when using randomized SVD to determine the eigenvalue decomposition for the low-rank (high-fidelity) method, and our bi-fidelity approximation, as a function of  $r$ . The bi-fidelity approximation requires  $2\times$  fewer FLOPs than the low-rank approach. Figure 5.8 (b) provides the comparison when using randomized Lanczos to determine the eigenvalue decomposi-

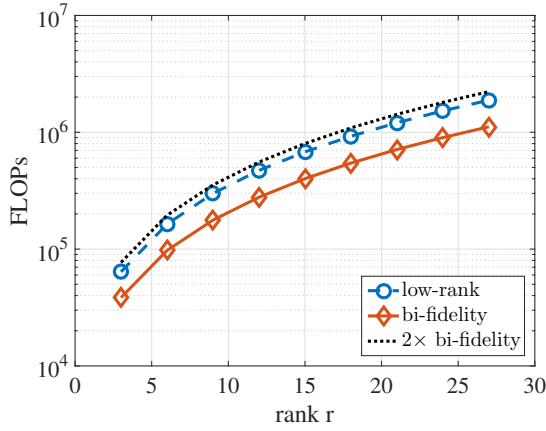
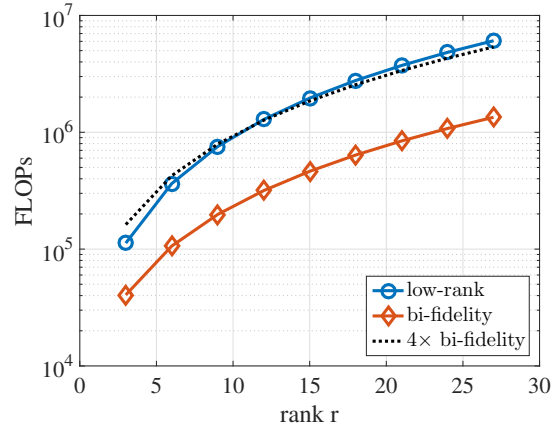
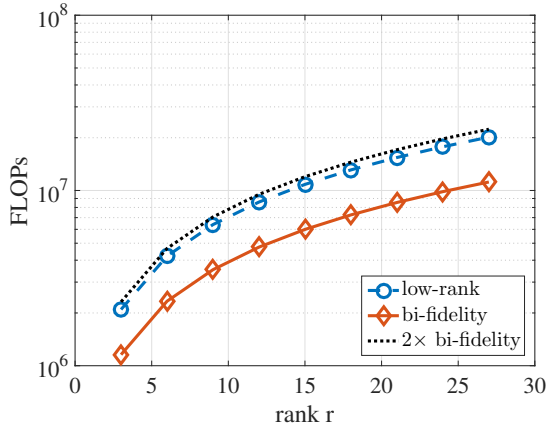
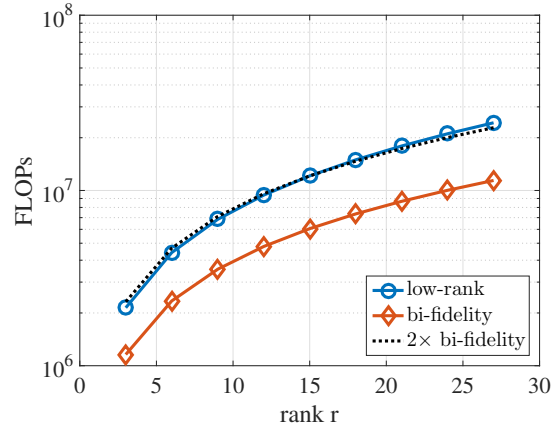
(a) randomized SVD,  $n_t = 1$ (b) randomized Lanczos,  $n_t = 1$ (c) randomized SVD,  $n_t = 45$ (d) randomized Lanczos,  $n_t = 45$ 

Figure 5.8: Cost comparisons between the low-rank (high-fidelity) method and our bi-fidelity approximation, as a function of rank  $r$ , for (a), (c) randomized SVD and (b), (d) randomized Lanczos for eigenvalue decomposition calculations. These estimates are based on FLOP decompositions. (a)-(b) compare the cost when using only  $n_t = 1$  time steps and (c)-(d) when using  $n_t = 45$  time steps. Dashed lines indicate the cost difference of the two methods.

tion for the low-rank (high-fidelity) method, and our bi-fidelity approximation, as a function of  $r$ . Here the bi-fidelity approximation requires  $4\times$  fewer FLOPs than the low-rank approach for all  $r$ . Figure 5.8 (c) and (d) provide the FLOP comparisons when using  $n_t = 45$  time steps. In both Figure 5.8 (c) and (d), both low-rank (high-fidelity) methods require  $2\times$  more FLOPs than the bi-fidelity approximation. From these results, we see that when the cost of the solver is closer to the number of degrees of freedom of the system, i.e.,  $N$ , the cost improvement of the bi-fidelity approximation is better when using randomized Lanczos. When the difference between  $\mathcal{C}_F(N)$  and  $N$  is larger, then the cost difference between performing the low-rank approach and bi-fidelity approach is primarily attributed to the number of high-fidelity models solves required. From the previous error estimates, if we only require ranks  $3 \leq r \leq 5$ , we will see comparable errors between the low-rank and bi-fidelity approximations, with a reduced cost.

### 5.5.2 Test Case # 2: $m_0 = 49$

For the second test case, we consider the scenario where we have a  $7 \times 7$  grid of sensors uniformly distributed over the spatial domain ( $m_0 = 49$ ) and use three time snapshots of data, resulting in  $m = 147$ . Thus, the rank of this system is larger than in test case #1. For the high-fidelity mapping, we consider a fine grid resolution with mesh size  $33 \times 33$ , and for the low-fidelity mapping, we consider a coarse grid resolution with mesh size  $17 \times 17$ , both of which are uniform. Thus  $N = 961$  and  $n = 225$ . Figure 5.9 provides non-boundary data for (a) the fine grid and (b) the coarse grid meshes with red “x” markers indicating the sensor points, or more precisely, the points at which the data is extracted. The surface magnitudes display the final solution of the PDE in (5.25) for reference.

Similar to the first example, we again treat the noise of these observations  $\mathbf{e}$  to be realizations of  $\mathcal{N}(\mathbf{0}, \mathbf{\Gamma}_{\text{noise}})$ , where the noise covariance is a diagonal matrix with entries  $\sigma_{\text{noise}}^2 \mathbf{A}\mathbf{u}_0$  with  $\sigma_{\text{noise}} = 0.03$ . For the prior covariance, we define  $\mathbf{\Gamma}_{\text{prior}}^{-1} = \alpha \mathbf{L}\mathbf{L}^T$ , where  $\mathbf{L}$  is the Laplacian operator. Here  $\alpha$  will be selected independently of the first example. Figure 5.10 provides the L-curve which is the plot of the error term of (5.8) given as  $\|\mathbf{\Gamma}_{\text{noise}}^{-1/2}(\mathbf{A}\mathbf{u}_\alpha - \mathbf{y})\|$  against the magnitude of the regularized

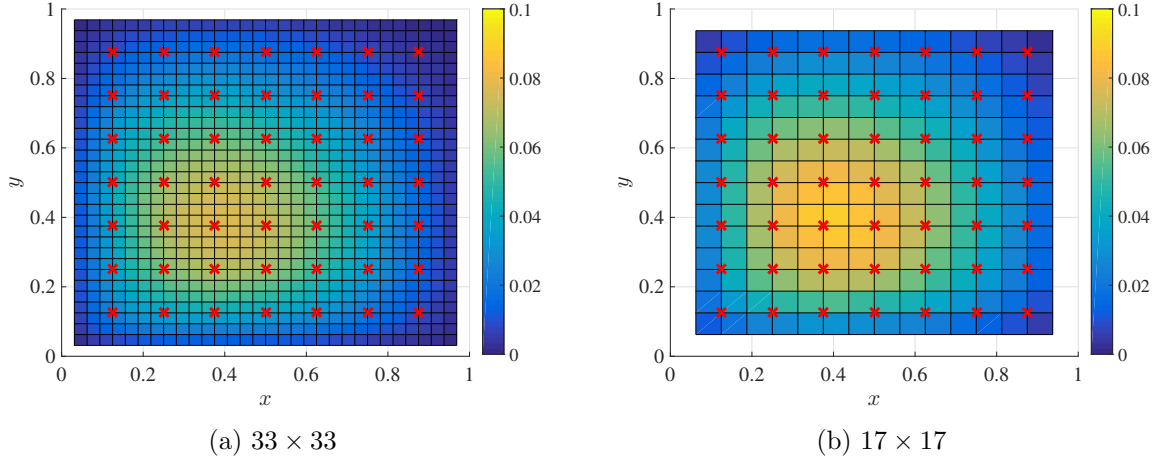


Figure 5.9: Numerical solutions (excluding boundary) to the PDE in (5.25) for (a) the fine grid mesh and (b) the coarse grid mesh. Red “x” markers indicate where observations are extracted.

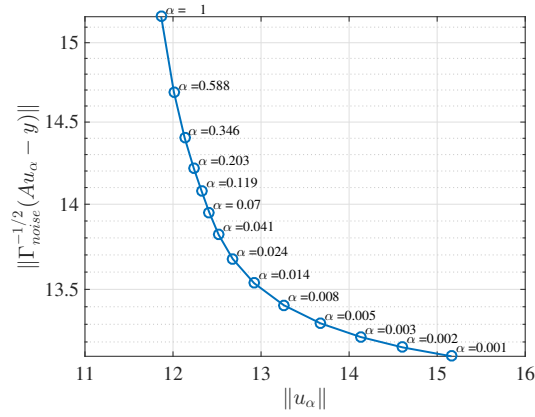


Figure 5.10: L-curve comparing the error term and regularization term of (5.8). Optimal point of  $\alpha = 0.014$  selected as balances the two terms.

solution given as  $\|\mathbf{u}_\alpha\|$ . This suggests that  $\alpha = 0.014$  is good choice for the regularization parameter.

As in Section 5.5.1, we select two values for the regularization term:  $\alpha = 0.014$  as per the L-curve, and  $\alpha = 200$ . Figure 5.11 provides the decay of eigenvalues for the high-fidelity and low-fidelity versions of the preconditioned data misfit of the Hessian for (a)  $\alpha = 0.014$  and (b)  $\alpha = 200$ . Similar to the previous set of results, the two figures display similar decay rates of the eigenvalues. The decay of the eigenvalues in Figure 5.11 indicates that a low-rank approximation to the data misfit of the Hessian will form a good approximation, where we will need rank  $r$  much less than  $m$ .

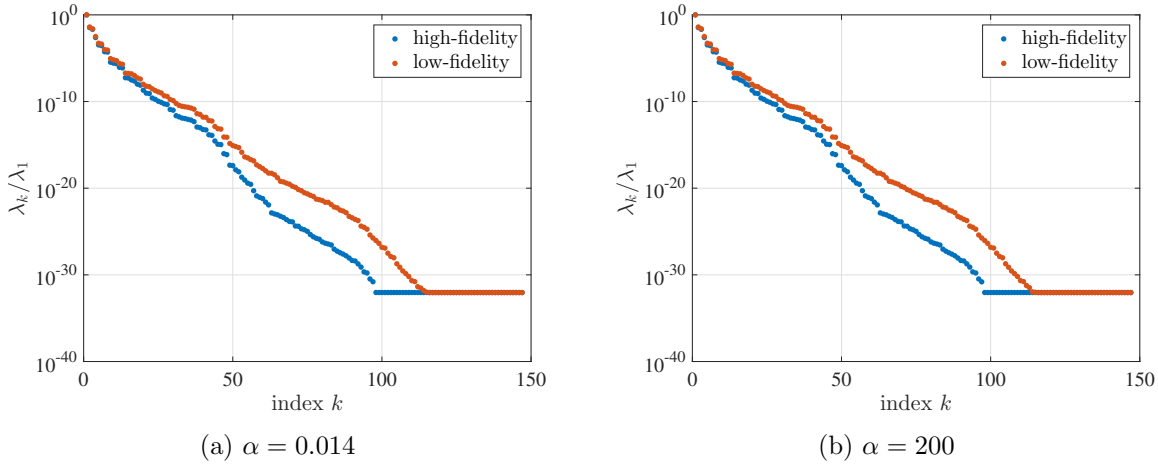


Figure 5.11: Decay of the eigenvalues of the data misfit of the Hessian for regularization parameter values (a)  $\alpha = 0.014$  and (b)  $\alpha = 200$ . High-fidelity refers the eigenvalues  $\mathbf{\Lambda}_r$ , as in (5.10), and low-fidelity refers to eigenvalues  $\tilde{\mathbf{\Lambda}}_r$ , as in (5.14).

The main result we consider for this example is the decay of error, as in (5.28), as a function of rank  $r$ . Figure 5.12 provides this error, where the left plots compare the low-rank (high-fidelity) estimate and the bi-fidelity estimate when using direct SVD, and the right plots compare the low-rank (high-fidelity) estimate and the bi-fidelity estimate when using randomized methods. For these results we look at ranks  $r \leq 60$ , so that we can better observe the behavior at smaller ranks. For larger ranks, the calculated errors maintain constant orders of magnitude. Figure 5.12 (a) and (b) provide these results for when  $\alpha = 0.014$ , as per the L-curve in Figure 5.10. In Figure 5.12 (a) we see an agreement in the error magnitudes for the low-rank and bi-fidelity approximations

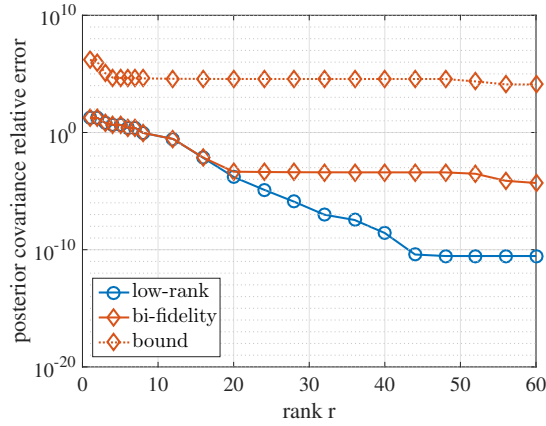
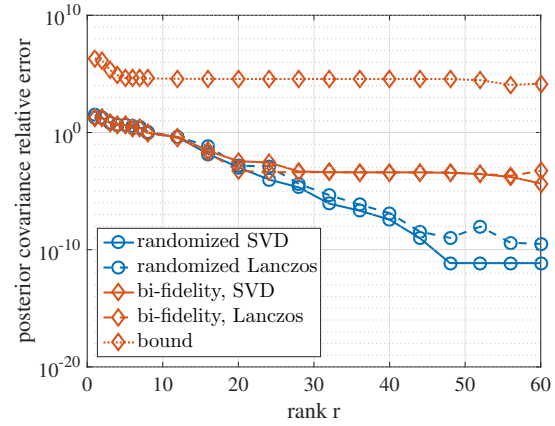
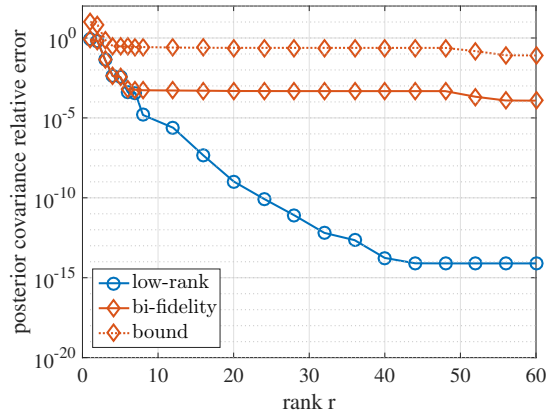
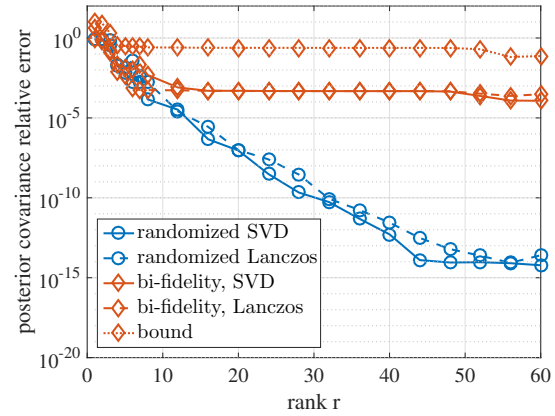
(a) SVD,  $\alpha = 0.014$ (b) randomized,  $\alpha = 0.014$ (c) SVD,  $\alpha = 200$ (d) randomized,  $\alpha = 200$ 

Figure 5.12: Relative error of the low-rank and low-rank, bi-fidelity approximations as in (5.28) for the scenarios (a)  $\alpha = 0.014$  and (b)  $\alpha = 200$ . Plots on the left correspond to the use of direct SVD to calculate the eigenvalue decompositions, and plots on the right correspond to using randomized approaches.

until rank  $r = 20$ . At this point, the low-rank approximation has a relative error of 0.0001 and the bi-fidelity approximation has a relative error of 0.0004. Thus if we require three to four digits of accuracy for this estimate, we can apply the bi-fidelity approximation as a surrogate. For this example, we consider the error bound as in (5.29). The error bound is significantly larger than the calculated error, with a relative error bound of about  $10^4$ . While it is not a tight bound, it does provide a bound on the error. Figure 5.12 (b) presents the calculated relative errors from the randomized approaches, where we see that the errors of the low-rank methods and bi-fidelity approach are similar until around rank  $r = 20$ . At rank  $r = 20$  randomized SVD and Lanczos have an error of 0.001, bi-fidelity with randomized SVD has an error of 0.003, and bi-fidelity with randomized Lanczos as an error of 0.0005. The error bound calculated via randomized Lanczos has an error of  $10^4$ . From these two figures, we observe that the direct methods are slightly more accurate for all low-rank methods, as well as for the error bound calculation, and that the bi-fidelity approximation is as accurate as the low-rank (high-fidelity) for smaller ranks.

To investigate the effect of the regularization parameter  $\alpha$  on the error, we consider an increased value of  $\alpha = 200$  in Figure 5.12 (c) and (d). Figure 5.12 (c) provides comparisons between the low-rank (high-fidelity) approximation and bi-fidelity approximation when using direct SVD to calculate the eigenvalue decomposition. Here the relative errors show agreement until rank  $r = 7$ , where both the low-rank (high-fidelity) method and the bi-fidelity approach have an error of 0.0005. For rank  $r \leq 4$ , the error bound is within an order of magnitude of the calculated bi-fidelity error. For larger ranks, the bound is about two orders of magnitude larger than the calculated error. Figure 5.12 (d) provides the error calculations when using randomized approaches. Similarly, the errors have close agreement until rank  $r = 8$ , with randomized SVD having an error of 0.0001, randomized Lanczos an error of 0.001, bi-fidelity approximation with randomized SVD an error of 0.005, and bi-fidelity approximation with randomized Lanczos an error of 0.0005. At rank  $r = 8$  the error bound is 0.25. For larger values of  $r$ , the low-rank (high-fidelity) methods continue to improve until about rank  $r = 50$ . After this point the errors of randomized SVD remain constant, and the errors randomized Lanczos become less accurate. From these results, we see again, that the



bi-fidelity approximation is as accurate as the low-rank (high-fidelity) method for smaller values of  $r$ .

With regards to the effect of the regularization parameter  $\alpha$ , we observe that larger values of  $\alpha$  result in a faster decay of the error with respect to the approximation rank for all methods. For the bi-fidelity approximation, in particular, we observe a similar error convergence for both values of  $\alpha$ , where larger values of  $\alpha$  result in a faster initial decay of the error. The ability of the error bound to provide an error estimation is significantly improved with an increase in  $\alpha$ .

Next we compare how high-fidelity, low-fidelity, and bi-fidelity approaches compare for posterior mean calculation. In particular, we are interested in identifying the coordinate of the maximum point of  $\mathbf{u}_0$ , (0.67, 0.67). Figure 5.13 provides this comparison as a function of the approximation rank  $r$  for both values of  $\alpha$ , but when using only direct SVD (as opposed to randomized methods). This is to show that we do in fact require going to the fine grid resolution of the high-fidelity model, as the coarse grid resolution of the low-fidelity model isn't resolved enough to yield similar accuracies. Figure 5.13 (a) provides the error comparison for  $\alpha = 0.014$ , where the high-fidelity and bi-fidelity errors agree at rank  $r = 12$ , with an error of 0.02, and the low-fidelity error is 0.08. Figure 5.13 (b) provides the comparison when  $\alpha = 200$ . For ranks  $r \geq 2$ , the high-fidelity and bi-fidelity approximations maintain an error of 0.02, while the low-fidelity approximation has an error of about 0.05. From both figures, the high-fidelity and bi-fidelity approaches outperform the low-fidelity approach.

Finally, we compare the costs of performing the different low-rank and bi-fidelity estimates of the posterior covariance, where FLOP estimates are calculated as discussed in Section 5.4.2. Figure 5.14 provides the comparisons of the FLOP estimates as a function of rank  $r$  for the randomized methods. Figure 5.14 (a) and (b) show the FLOP counts for  $n_t = 1$  time step. Little cost improvement is obtained when using randomized SVD. For the randomized Lanczos approach, the bi-fidelity approximation has a cost improvement of at least  $2\times$ . Figure 5.14 (c) and (d) show the FLOP counts for  $n_t = 45$  time steps, where both randomized approaches provide a  $2\times$  improvement in the bi-fidelity cost. Compared to the first example of Section 5.5.1, the cost reduction is smaller.

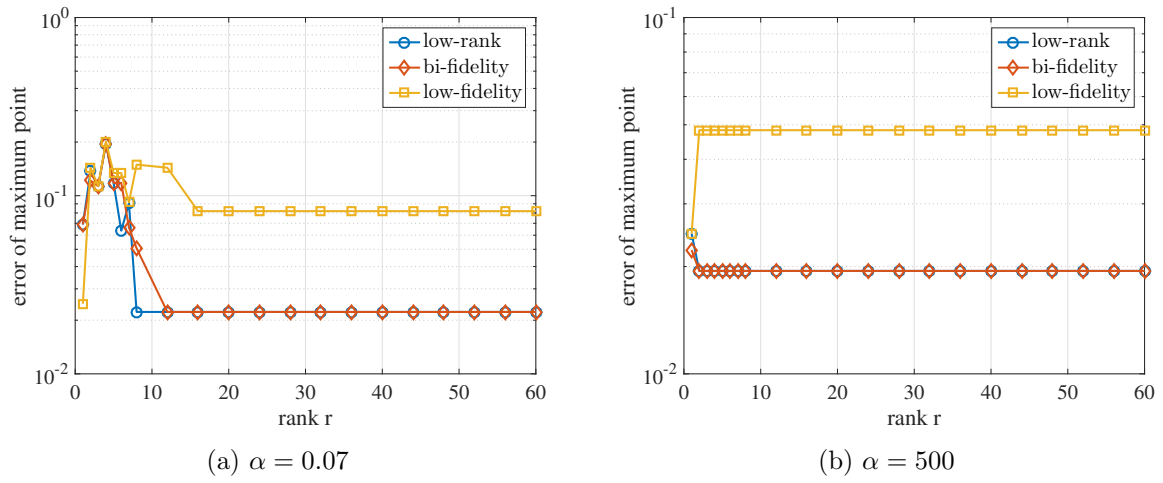


Figure 5.13: Euclidean distance between the true maximum point coordinate  $(0.67, 0.67)$  and the maximum point coordinate as calculated from the posterior mean (via MAP estimation) in (5.8). Three methods are compared: low-rank (high-fidelity), new bi-fidelity, and low-rank, low-fidelity approaches, as a function of rank  $r$ .

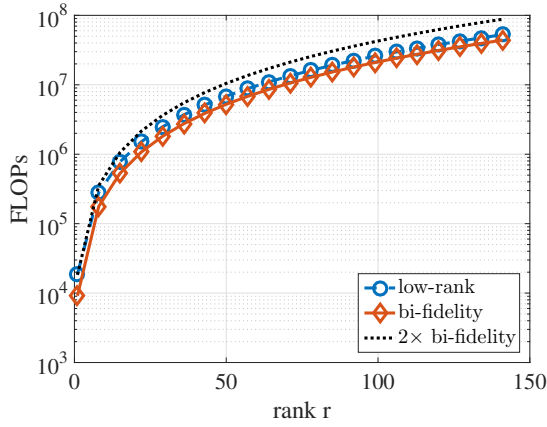
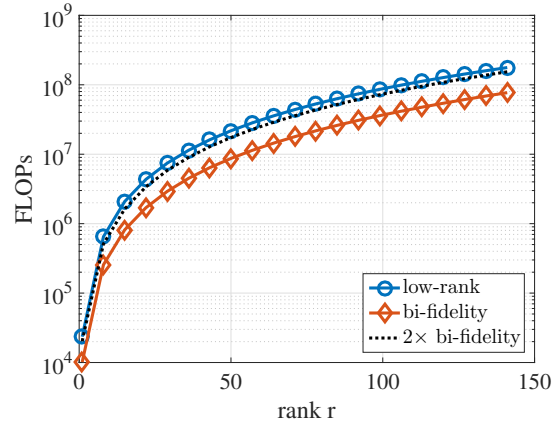
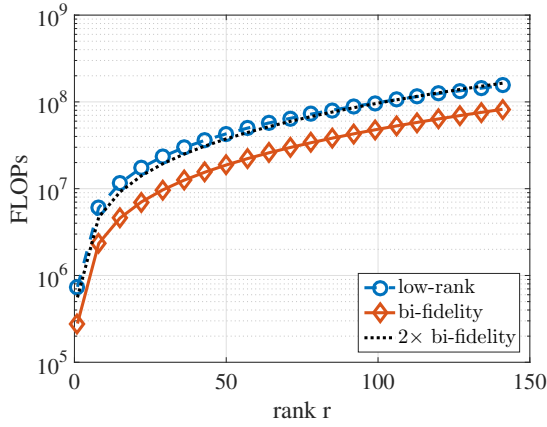
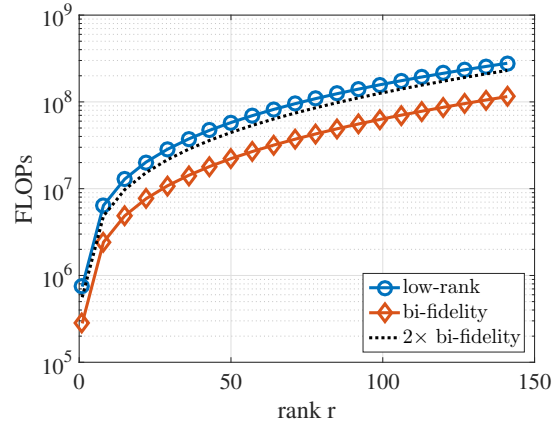
(a) randomized SVD,  $n_t = 1$ (b) randomized Lanczos,  $n_t = 1$ (c) randomized SVD,  $n_t = 45$ (d) randomized Lanczos,  $n_t = 45$ 

Figure 5.14: Cost comparisons between the low-rank (high-fidelity) method and our bi-fidelity approximation, as a function of rank  $r$ , for (a), (c) randomized SVD and (b), (d) randomized Lanczos for eigenvalue decomposition calculations. These estimates are based on FLOP decompositions. (a)-(b) compare the cost when using only  $n_t = 1$  time steps and (c)-(d) when using  $n_t = 45$  time steps. Dashed lines indicate the cost difference of the two methods.

This is primarily due to the fact that the cost difference of the low-fidelity and high-fidelity models is smaller. In the previous example the high-fidelity model was  $16\times$  more expensive than the low-fidelity model, while here the cost difference is  $4\times$ .

## 5.6 Conclusions and Future Work

In this work we present a novel low-rank, bi-fidelity approach for the estimation of the posterior covariance as determined with MAP estimation. This formulation, which employs ideas from the Nyström approximation, estimates the high-fidelity eigenvectors using the eigenvectors and eigenvalues derived from a low-fidelity mapping. The estimated eigenvectors and low-fidelity eigenvalues may then be used to form a low-rank approximation to the posterior covariance.

In addition to guaranteed cost reduction, theoretical error analysis of this work provides an error bound, and in some scenarios an error estimate, to this bi-fidelity approximation. In particular, if we assume a small error in the linear mapping between low-fidelity and high-fidelity maps, we may show that the error of the bi-fidelity approximation is small as well. Furthermore, when the rank  $r$  approximations of both the low-fidelity and high-fidelity operators map to the same range, this bi-fidelity approximation will have the same accuracy as the low-rank, high-fidelity approximation to the posterior covariance.

The numerical examples of this work highlight the effectiveness of this approximation for linear Bayesian inference. Specifically, the bi-fidelity approximation can have two to four digits of accuracy, allowing for similar accuracies as low-rank, high-fidelity methods up to a certain rank, but with a significantly reduced cost.

Error results from numerical examples indicate a significant impact of the magnitude of the regularization parameter  $\alpha$  on both the accuracy of the bi-fidelity approximation and effectiveness of the error bound from Theorem 5.4.1. In particular, there is a trade-off between the approximation accuracy and the error bound’s usefulness as an error estimate. Future work will seek to explore this connection, and identify additional theoretical analysis to obtain a more effective error bound for the bi-fidelity approximation in these varying scenarios.

## Bibliography

- [1] Exascale Computing Engineering Center. Predictive Science Academic Alliance Program (PSAAP) II, Stanford University, 2017.
- [2] Mira, Argonne Leadership Computing Facility, 2017.
- [3] Titan Cray XK7, Oak Ridge Leadership Computing Facility, 2017.
- [4] S.S. Adavani and G. Biros. Multigrid algorithms for inverse problems with linear parabolic PDE constraints. SIAM Journal on Scientific Computing, 31(1):369–397, 2008.
- [5] B. O. Almroth, P. Stern, and F. A. Brogan. Automatic choice of global shape functions in structural analysis. AIAA Journal, 16(5):525–528, 1978.
- [6] P. Arbenz and D. Kressner. Lecture notes on solving large scale eigenvalue problems. D-MATH, EHT Zurich, 2, 2012.
- [7] S.R. Arridge, J.P. Kaipio, V. Kolehmainen, M. Schweiger, E. Somersalo, T. Tarvainen, and M. Vauhkonen. Approximation errors and model reduction with an application in optical diffusion tomography. Inverse Problems, 22(1):175, 2006.
- [8] S. Asmussen and P. Glynn. Stochastic simulation: algorithms and analysis, volume 57. Springer Science & Business Media, 2007.
- [9] M. Bachmayr and A. Cohen. Kolmogorov widths and low-rank approximations of parametric elliptic PDEs. Mathematics of Computation, 86(304):701–724, 2017.
- [10] S. Balachandar and J. K. Eaton. Turbulent dispersed multiphase flow. Annual Review of Fluid Mechanics, 42:111–133, 2010.
- [11] A. Barth, C. Schwab, and N. Zollinger. Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients. Numerische Mathematik, 119(1):123–161, 2011.
- [12] P. Benner, Y. Qiu, and M. Stoll. Low-rank computation of posterior covariance matrices in bayesian inverse problems. arXiv preprint arXiv:1703.05638, 2017.
- [13] G. Biros and G. Dogan. A multilevel algorithm for inverse problems with elliptic PDE constraints. Inverse Problems, 24(3):034010, 2008.
- [14] J. L. Bouchot, H. Rauhut, and Christoph C. Schwab. Multi-level compressed sensing Petrov-Galerkin discretization of high-dimensional parametric PDEs. arXiv preprint arXiv:1701.01671, 2017.

- [15] A. Brandt. Multi-level adaptive solutions to boundary-value problems. Mathematics of computation, 31(138):333–390, 1977.
- [16] H. Brenner. Effect of finite boundaries on the Stokes resistance of an arbitrary particle. Journal of Fluid Mechanics, 12:35–48, 1962.
- [17] W. L. Briggs, V. E. Henson, and S. F. McCormick. A multigrid tutorial. 2000.
- [18] T. Bui-Thanh, C. Burstedde, O. Ghattas, J. Martin, G. Stadler, and L.C. Wilcox. Extreme-scale UQ for Bayesian inverse problems governed by PDEs. In Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis, page 3. IEEE Computer Society Press, 2012.
- [19] T. Bui-Thanh, O. Ghattas, J. Martin, and G. Stadler. A computational framework for infinite-dimensional Bayesian inverse problems part i: The linearized case, with application to global seismic inversion. SIAM Journal on Scientific Computing, 35(6):A2494–A2523, 2013.
- [20] T. Bui-Thanh, K. Willcox, and O. Ghattas. Model reduction for large-scale systems with high-dimensional parametric input space. SIAM Journal on Scientific Computing, 30(6):3270–3288, 2008.
- [21] R. Butler, T.J. Dodwell, T. Kim, S. Kynaston, R. Scheichl, R.T. Haftka, and N.H. Kim. Uncertainty quantification of composite structures with defects using multilevel Monte Carlo simulations. In AIAA SciTech Conf, 2015.
- [22] M. Caporali, F. Tampieri, F. Trombetti, and O. Vittori. Transfer of particles in nonisotropic air turbulence. Journal of Atmospheric Sciences, 32:565–568, 1975.
- [23] H. Cheng, Z. Gimbutas, P.G. Martinsson, and V. Rokhlin. On the compression of low rank matrices. SIAM Journal on Scientific Computing, 26(4):1389–1404, 2005.
- [24] K.A. Cliffe, M.B. Giles, R. Scheichl, and A.L. Teckentrup. Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients. Computing and Visualization in Science, 14(1):3–15, 2011.
- [25] P. G. Constantine, M. Eldred, and E. Phipps. Sparse pseudospectral approximation method. Computer Methods in Applied Mechanics and Engineering, 229:1–12, 2012.
- [26] M.S. Eldred D.M. and Dunlavy. Formulations for surrogate-based optimization with data fit, multifidelity, and reduced-order models. In Proceedings of the 11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, number AIAA-2006-7117, Portsmouth, VA, volume 199, 2006.
- [27] A. Doostan, G. Geraci, and G. Iaccarino. A bi-fidelity approach for uncertainty quantification of heat transfer in a rectangular ribbed channel. In ASME Turbo Expo 2016: Turbomachinery Technical Conference and Exposition, pages V02CT45A031–V02CT45A031. American Society of Mechanical Engineers, 2016.
- [28] A. Doostan, R. Ghanem, and J. Red-Horse. Stochastic model reduction for chaos representations. Computer Methods in Applied Mechanics and Engineering, 196(37-40):3951–3966, 2007.

- [29] A. Doostan and H. Owhadi. A non-adapted sparse approximation of PDEs with stochastic inputs. Journal of Computational Physics, 230:3015–3034, 2011.
- [30] E.L. Dyer, T.A. Goldstein, R. Patel, K.P. Kording, and R.G. Baraniuk. Self-expressive decompositions for matrix approximation and clustering. arXiv preprint arXiv:1505.00824, 2015.
- [31] Y. Efendiev, T. Hou, and W. Luo. Preconditioning markov chain Monte Carlo simulations using coarse-scale models. SIAM Journal on Scientific Computing, 28(2):776–803, 2006.
- [32] M.S. Eldred. Recent advances in non-intrusive polynomial chaos and stochastic collocation methods for uncertainty analysis and design. AIAA Paper, 2274(2009):37, 2009.
- [33] E. Elhamifar and R. Vidal. Sparse subspace clustering. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, pages 2790–2797. IEEE, 2009.
- [34] M. Emsermann and B. Simon. Improving simulation efficiency with quasi control variates. 2002.
- [35] M. Esmaily, L. Jofre, A. Mani, and G. Iaccarino. A scalable geometric multigrid solver for nonsymmetric elliptic systems with application to variable-density flows. Journal of Computational Physics, 357:142–158, 2018.
- [36] H. R. Fairbanks, A. Doostan, C. Ketelsen, and G. Iaccarino. A low-rank control variate for multilevel Monte Carlo simulation of high-dimensional uncertain systems. Journal of Computational Physics, 341:121–139, 2017.
- [37] H. R. Fairbanks, L. Jofre, G. Geraci, G. Iaccarino, and A. Doostan. A bi-fidelity approximation for uncertainty quantification and sensitivity analysis of irradiated particle-laden turbulence. In prep.
- [38] E. Farbar, I. D. Boyd, and M. Esmaily-Moghadam. Monte Carlo modeling of radiative heat transfer in particle-laden flow. Journal of Quantitative Spectroscopy and Radiative Transfer, 184:146–160, 2017.
- [39] M. G. Fernández-Godino, C. Park, N. H. Kim, and R. T. Haftka. Review of multi-fidelity models. arXiv preprint arXiv:1609.07196, 2016.
- [40] H.P. Flath, L.C. Wilcox, V. Akcelik, J. Hill, B. van Bloemen Waanders, and O. Ghattas. Fast algorithms for bayesian uncertainty quantification in large-scale linear inverse problems based on low-rank partial hessian approximations. SIAM Journal on Scientific Computing, 33(1):407–432, 2011.
- [41] A.I.J. Forrester, A. Sóbester, and A.J. Keane. Multi-fidelity optimization via surrogate modelling. In Proceedings of the royal society of london a: mathematical, physical and engineering sciences, volume 463, pages 3251–3269. The Royal Society, 2007.
- [42] C. Fowlkes, S. Belongie, F. Chung, and J. Malik. Spectral grouping using the nystrom method. IEEE transactions on pattern analysis and machine intelligence, 26(2):214–225, 2004.
- [43] A. Frankel and G. Iaccarino. Efficient control variates for uncertainty quantification of radiation transport. Journal of Quantitative Spectroscopy and Radiative Transfer, 189:398–406, 2017.

- [44] A. Frankel, H. Pouransari, F. Coletti, and A. Mani. Settling of heated particles in homogeneous turbulence. Journal of Fluid Mechanics, 792:869–893, 2016.
- [45] S. Gangadharan, R. Haftka, and Y. Fiocca. Variable-complexity-modeling structural optimization using response surface methodology. In 36th Structures, Structural Dynamics and Materials Conference, page 1164, 1995.
- [46] G. Geraci, M. Eldred, and G. Iaccarino. A multifidelity control variate approach for the multilevel Monte Carlo technique. Center for Turbulence Research Annual Research Briefs, 2015.
- [47] R. Ghanem, G. Saad, and A. Doostan. Efficient solution of stochastic systems: Application to the embankment dam problem. Structural Safety, 29(3):238–251, 2007.
- [48] R. Ghanem and P. Spanos. Stochastic Finite Elements: A Spectral Approach. Dover, 2002.
- [49] M. B. Giles. Multilevel Monte Carlo path simulation. Operations Research, 56(3):607–617, 2008.
- [50] M. B. Giles. Multilevel Monte Carlo methods. In Monte Carlo and Quasi-Monte Carlo Methods 2012, pages 83–103. Springer, 2013.
- [51] G. H. Golub and C. F. Van Loan. Matrix computations (3rd ed.). Johns Hopkins University Press, Baltimore, MD, USA, 1996.
- [52] L. Le Gratiet and C. Cannamela. Cokriging-based sequential design strategies using fast cross-validation techniques for multi-fidelity computer codes. Technometrics, 57:418–427, 2015.
- [53] L. Le Gratiet and J. Garnier. Recursive co-kriging model for design of computer experiments with multiple levels of fidelity. International Journal for Uncertainty Quantification, 4(5), 2014.
- [54] M. Gu and S.C. Eisenstat. Efficient algorithms for computing a strong rank-revealing QR factorization. SIAM Journal on Scientific Computing, 17(4):848–869, 1996.
- [55] M. Hadigol and A. Doostan. Least squares polynomial chaos expansion: A review of sampling strategies. Computer Methods in Applied Mechanics and Engineering, 332:382–407, 2018.
- [56] 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, 300:507–524, 2015.
- [57] N. Halko, P.G. Martinsson, and J.A. Tropp. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. SIAM review, 53(2):217–288, 2011.
- [58] J.M. Hammersley and D.C. Handscomb. Monte Carlo methods, 1964.
- [59] J. Hampton and A. Doostan. Coherence motivated sampling and convergence analysis of least squares polynomial chaos regression. Computer Methods in Applied Mechanics and Engineering, 290:73–97, 2015.



- [60] J. Hampton and A. Doostan. Compressive sampling of polynomial chaos expansions: Convergence analysis and sampling strategies. Journal of Computational Physics, 280:363–386, 2015.
- [61] J. Hampton and A. Doostan. Basis adaptive sample efficient polynomial chaos (base-pc). arXiv preprint arXiv:1702.01185, 2017.
- [62] J. Hampton, H. Fairbanks, A. Narayan, and A. Doostan. Practical error bounds for a non-intrusive bi-fidelity approach to parametric/stochastic model reduction. Journal of Computational Physics, accepted and in press.
- [63] P. C. Hansen. Analysis of discrete ill-posed problems by means of the L-curve. SIAM review, 34(4):561–580, 1992.
- [64] P. C. Hansen. Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion, volume 4. Siam, 2005.
- [65] S. Heinrich. Multilevel Monte Carlo methods. In Large-scale scientific computing, pages 58–67. Springer, 2001.
- [66] C. K. Ho. Advances in central receivers for concentrating solar applications. Solar Energy, 152:38–56, 2017.
- [67] S. Hosder, R.W. 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.
- [68] J. Jiang, Y. Chen, and A. Narayan. Offline-enhanced reduced basis method through adaptive construction of the surrogate training set. Journal of Scientific Computing, 73(2-3):853–875, 2017.
- [69] L. Jiang and N. Ou. Multiscale model reduction method for bayesian inverse problems of subsurface flow. Journal of Computational and Applied Mathematics, 319:188–209, 2017.
- [70] L. Jofre, S. P. Domino, and G. Iaccarino. A framework for characterizing structural uncertainty in large-eddy simulation closures. Flow Turbulence and Combustion, 100(2):341–363, 2018.
- [71] L. Jofre, G. Geraci, H. R. Fairbanks, A. Doostan, and G. Iaccarino. Multi-fidelity uncertainty quantification of irradiated particle-laden turbulence. Technical Report Annual Research Brief, Center for Turbulence Research, Stanford University, 2017.
- [72] M.C. Kennedy and A. O’Hagan. Predicting the output from a complex computer code when fast approximations are available. Biometrika, 87(1):1–13, 2000.
- [73] W. Kleiber, S. R. Sain, M. J. Heaton, M. Wiltberger, C. S. Reese, D. Bingham, et al. Parameter tuning for a multi-fidelity dynamical model of the magnetosphere. The Annals of Applied Statistics, 7:1286–1310, 2013.
- [74] D. Kumar, M. Raisee, and C. Lacor. An efficient non-intrusive reduced basis model for high dimensional stochastic problems in cfd. Computers & Fluids, 138:67–82, 2016.

- [75] J. C. Lasheras and E. J. Hopfinger. Liquid jet instability and atomization in a coaxial gas stream. Annual Review of Fluid Mechanics, 32:275–308, 2000.
- [76] J. Laurenceau and P. Sagaut. Building efficient response surfaces of aerodynamic functions with kriging and cokriging, 2008.
- [77] E. Liberty, F. Woolfe, P. G. Martinsson, V. Rokhlin, and M. Tygert. Randomized algorithms for the low-rank approximation of matrices. Proceedings of the National Academy of Sciences, 104(51):20167–20172, 2007.
- [78] C. Lieberman, K. Fidkowski, K. Willcox, and B. van Bloemen Waanders. Hessian-based model reduction: large-scale inversion and prediction. International Journal for Numerical Methods in Fluids, 71(2):135–150, 2013.
- [79] A. Logg, K. A. Mardal, and G. N. Wells. Automated Solution of Differential Equations by the Finite Element Method. Springer, 2012.
- [80] M. Mahoney and P. Drineas. Cur matrix decompositions for improved data analysis. Proceedings of the National Academy of Sciences, 106:697–702, 2009.
- [81] O. Le Maitre, M. Reagan, H. Najm, R. Ghanem, and O. Knio. A stochastic projection method for fluid flow. ii: Random process. J. Comp. Phys., 181:9–44, 2002.
- [82] O. P. Le Maître and O. Knio. Spectral Methods for Uncertainty Quantification with Applications to Computational Fluid Dynamics. Springer, 2010.
- [83] P. G. Martinsson, G. Quintana Ortí, N. Heavner, and R. van de Geijn. Householder qr factorization with randomization for column pivoting (hqrrp). SIAM Journal on Scientific Computing, 39(2):C96–C115, 2017.
- [84] P.G. Martinsson, V. Rokhlin, and M. Tygert. A randomized algorithm for the decomposition of matrices. Applied and Computational Harmonic Analysis, 30(1):47–68, 2011.
- [85] Y. M. Marzouk and H. N. Najm. Dimensionality reduction and polynomial chaos acceleration of bayesian inference in inverse problems. J. Comput. Phys., 228:1862–1902, 2009.
- [86] Y. M. Marzouk and D. Xiu. A stochastic collocation approach to bayesian inference in inverse problems. Communications In Computational Physics, 6(4):826–847, 2009.
- [87] B. Mason, R. Haftka, and E. Johnson. Analysis and design of composite curved channel frames. In 5th Symposium on Multidisciplinary Analysis and Optimization, page 4364, 1994.
- [88] L. Mathelin and M.Y. Hussaini. A stochastic collocation algorithm for uncertainty analysis. Technical Report NAS 1.26:212153; NASA/CR-2003-212153, NASA Langley Research Center, 2003.
- [89] M. R. Maxey and J. J. Riley. Equation of motion for a small rigid sphere in a nonuniform flow. Physics of Fluids, 26:883–889, 1983.
- [90] J. L. Mueller and S. Siltanen. Linear and nonlinear inverse problems with practical applications, volume 10. Siam, 2012.

- [91] A. Narayan, C. Gittelsohn, and D. Xiu. A stochastic collocation algorithm with multifidelity models. SIAM Journal on Scientific Computing, 36(2):A495–A521, 2014.
- [92] L.W.-T. Ng and M.S. Eldred. Multifidelity uncertainty quantification using nonintrusive polynomial chaos and stochastic collocation. In Proceedings of the 14th AIAA Non-Deterministic Approaches Conference, number AIAA-2012-1852, Honolulu, HI, volume 45, 2012.
- [93] F. Nobile and F. Tesei. A multi level Monte Carlo method with control variate for elliptic PDEs with log-normal coefficients. Stochastic Partial Differential Equations: Analysis and Computations, 3(3):398–444, 2015.
- [94] A.K. Noor and J.M. Peters. Reduced basis technique for nonlinear analysis of structures. AIAA Journal, 18(4):455–462, 1980.
- [95] A.S. Padrón, J.J. Alonso, and M.S. Eldred. Multi-fidelity methods in aerodynamic robust optimization. In 18th AIAA Non-Deterministic Approaches Conference, page 0680, 2016.
- [96] P.S. Palar, T. Tsuchiya, and G. Parks. Decomposition-based evolutionary aerodynamic robust optimization with multi-fidelity point collocation non-intrusive polynomial chaos. In 17th AIAA Non-Deterministic Approaches Conference, page 1377, 2015.
- [97] L. Parussini, D. Venturi, P. Perdikaris, and G. E. Karniadakis. Multi-fidelity gaussian process regression for prediction of random fields. Journal of Computational Physics, 336:36–50, 2017.
- [98] R. Pasupathy, B.W. Schmeiser, and M.R. Taaffe J. Wang. Control-variate estimation using estimated control means. IIE Transactions, 44(5):381–385, 2012.
- [99] B. Peherstorfer, T. Cui, Y. Marzouk, and K. Willcox. Multifidelity importance sampling. Computer Methods in Applied Mechanics and Engineering, 300:490–509, 2016.
- [100] J. Peng, J. Hampton, and A. Doostan. A weighted  $\ell_1$ -minimization approach for sparse polynomial chaos expansions. Journal of Computational Physics, 267:92 – 111, 2014.
- [101] P. Perdikaris, D. Venturi, and G. E. Karniadakis. Multifidelity information fusion algorithms for high-dimensional systems and massive data sets. SIAM Journal on Scientific Computing, 38:B521–B538, 2016.
- [102] P. Perdikaris, D. Venturi, J. O. Royset, and G. E. Karniadakis. Multi-fidelity modelling via recursive co-kriging and gaussian–markov random fields. Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 471, 2015.
- [103] D. J. Perry, R. M. Kirby, A. Narayan, and R. T. Whitaker. Allocation strategies for high fidelity models in the multifidelity regime. submitted, 2017.
- [104] J. S. Peterson. The reduced basis method for incompressible viscous flow calculations. SIAM Journal on Scientific and Statistical Computing, 10:777–786, 1989.
- [105] A. Pinkus. N-widths in approximation theory. Springer, 1985.
- [106] C. Prud’Homme, D. V. Rovas, K. Veroy, L. Machiels, Y. Maday, A. T. Patera, and G. Turinici. Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bound methods. Journal of Fluids Engineering, 124(1):70–80, 2002.

- [107] Z. Qian, C.C. Seepersad, V.R. Joseph, J. K. Allen, and C.F.J. Wu. Building surrogate models based on detailed and approximate simulations. Journal of Mechanical Design, 128(4):668–677, 2006.
- [108] P. Le Quéré. Accurate solutions to the square thermally driven cavity at high rayleigh number. Computers & Fluids, 20(1):29–41, 1991.
- [109] M. Rahmani, G. Geraci, G. Iaccarino, and A. Mani. Polydisperse particles in an irradiated turbulent gas-particle mixture. Center for Turbulence Research Annual Research Briefs, Stanford University, pages 27–41, 2015.
- [110] V. Raman and R. O. Fox. Modeling of fine-particle formation in turbulent flames. Annual Review of Fluid Mechanics, 48:159–190, 2016.
- [111] C. E. Rasmussen and C.K. Williams. Gaussian processes for machine learning, volume 1. MIT press Cambridge, 2006.
- [112] V. Rokhlin, A. Szlam, and M. Tygert. A randomized algorithm for principal component analysis. SIAM Journal on Matrix Analysis and Applications, 31(3):1100–1124, 2009.
- [113] S. T. Roweis. EM algorithms for PCA and SPCA. In Advances in neural information processing systems, pages 626–632, 1998.
- [114] G. Rozza, D. Huynh, and A. Patera. Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations. Archives of Computational Methods in Engineering, 15(3):229–275, 2008.
- [115] O. Rubio, E. Bravo, and J.R. Claeysen. Thermally driven cavity flow with neumann condition for the pressure. Applied numerical mathematics, 40(1):327–336, 2002.
- [116] J. Sacks, W.J. Welch, T.J. Mitchell, and H.P. Wynn. Design and analysis of computer experiments. Statistical science, pages 409–423, 1989.
- [117] A. K. Saibaba and P.K. Kitanidis. Efficient methods for large-scale linear inversion using a geostatistical approach. Water Resources Research, 48(5), 2012.
- [118] A. K. Saibaba and P.K. Kitanidis. Fast computation of uncertainty quantification measures in the geostatistical approach to solve inverse problems. Advances in Water Resources, 82:124–138, 2015.
- [119] G. Sardina, P. Schlatter, L. Brandt, and F. Picano. Wall accumulation and spatial localization in particle-laden wall flows. Journal of Fluid Mechanics, 699:50–78, 2012.
- [120] B. Schmeiser, M. Taaffe, and J. Wang. Biased control-variate estimation. IIE Transactions, 33(3):219–228, 2001.
- [121] R. A. Shaw. Particle-turbulence interactions in atmospheric clouds. Annual Review of Fluid Mechanics, 35:183–227, 2003.
- [122] R. W. Skinner, A. Doostan, E. L. Peters, J. A. Evans, and K. E. Jansen. An evaluation of bi-fidelity modeling efficiency on a general family of naca airfoils. In 35th AIAA Applied Aerodynamics Conference, page 3260, 2017.

- [123] A. Spantini, A. Solonen, T. Cui, J. Martin, L. Tenorio, and Y. Marzouk. Optimal low-rank approximations of bayesian linear inverse problems. SIAM Journal on Scientific Computing, 37(6):A2451–A2487, 2015.
- [124] A. Speight. A multilevel approach to control variates. Journal of Computational Finance, Forthcoming, 2009.
- [125] K. D. Squires and J. K. Eaton. Preferential concentration of particles by turbulence. Physics of Fluids, 3:1169–1178, 1991.
- [126] B. Sudret. Global sensitivity analysis using polynomial chaos expansions. Reliability Engineering & System Safety, 93(7):964–979, 2008.
- [127] W. Sutherland. LII. the viscosity of gases and molecular force. Philosophical Magazine, 36:507–531, 1893.
- [128] A.L. Teckentrup, R. Scheichl, M.B. Giles, and E. Ullmann. Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients. Numerische Mathematik, 125(3):569–600, 2013.
- [129] S. R. Tieszen. On the fluid mechanics of fires. Annual Review of Fluid Mechanics, 33:67–92, 2001.
- [130] F. Veron. Ocean spray. Annual Review of Fluid Mechanics, 47:507–538, 2015.
- [131] F. Vidal-Codina, N.C. Nguyen, M. B. Giles, and J. Peraire. A model and variance reduction method for computing statistical outputs of stochastic elliptic partial differential equations. Journal of Computational Physics, 297:700–720, 2015.
- [132] Q. Wang and K. D. Squires. Large eddy simulation of particle-laden turbulent channel flow. Physics of Fluids, 8:1207–1223, 1996.
- [133] R. C. Weast, M. J. Astle, and W. H. Beyer. Handbook of chemistry and physics. CRC Press, 1989.
- [134] C. K. Williams and M. Seeger. Using the Nyström method to speed up kernel machines. In Advances in neural information processing systems, pages 682–688, 2001.
- [135] D. Xiu and J.S. Hesthaven. High-order collocation methods for differential equations with random inputs. SIAM J. Sci. Comput., 27(3):1118–1139, 2005.
- [136] D. Xiu and G.M. Karniadakis. The Wiener-Askey polynomial chaos for stochastic differential equations. SIAM Journal on Scientific Computing, 24(2):619–644, 2002.
- [137] D. Xiu, Y. Kevrekidis, and R. Ghanem. An equation-free, multiscale approach to uncertainty quantification. Computing in Science and Engineering, 7(3):16–23, 2005.
- [138] F.-L. Yang and M. L. Hunt. Dynamics of particle-particle collisions in a viscous liquid. Physics of Fluids, 18:121506, 2006.
- [139] R. Zamansky, F. Coletti, M. Massot, and A. Mani. Radiation induces turbulence in particle-laden fluids. Physics of Fluids, 26:071701, 2014.

- [140] A. Zaytsev and E. Burnaev. Large scale variable fidelity surrogate modeling. Annals of Mathematics and Artificial Intelligence, 81(1-2):167–186, 2017.
- [141] K. Zhang and J.T. Kwok. Clustered nyström method for large scale manifold learning and dimension reduction. IEEE Transactions on Neural Networks, 21(10):1576–1587, 2010.
- [142] X. Zhu, A. Narayan, and D. Xiu. Computational aspects of stochastic collocation with multifidelity models. SIAM/ASA Journal on Uncertainty Quantification, 2(1):444–463, 2014.

## Appendix A

### Matrix Interpolative Decomposition

The following lemma from [84] outlines a number of key properties of the ID representation of  $\mathbf{U}_{\ell-1}$ , as seen in (4.15). For simplicity, we drop the subscripts of  $\mathbf{U}$ ,  $\mathbf{C}$ ,  $\mathbf{U}^c$ ,  $m$ , and  $N$ .

**Lemma A.0.1.** *(Lemma 3.1 of [84].) For any positive integer  $r$  with  $r \leq \min\{m, N\}$ , there exist a real  $r \times N$  matrix  $\mathbf{C}$ , and a real  $m \times r$  matrix  $\mathbf{U}^c$  whose columns constitute a subset of the columns of  $\mathbf{U}$ , such that*

- (1) *some subset of the columns of  $\mathbf{C}$  makes up the  $r \times r$  identity matrix,*
- (2) *no entry of  $\mathbf{C}$  has an absolute value greater than 1,*
- (3)  $\|\mathbf{C}\| \leq \sqrt{r(N-r)+1},$
- (4) *the least (that is, the  $r$ th greatest) singular value of  $\mathbf{C}$  is at least 1,*
- (5)  $\mathbf{U} = \mathbf{U}^c \mathbf{C}$ , *when  $r = m$  or  $r = N$ , and*
- (6)  $\|\mathbf{U} - \mathbf{U}^c \mathbf{C}\| \leq \sqrt{r(N-r)+1} \sigma_{r+1}$  *when  $r < \min\{m, N\}$ , where  $\sigma_{r+1}$  is the  $(r+1)$ st greatest singular value of  $\mathbf{U}$ .*

The construction of matrix ID relies primarily on the rank-revealing QR factorization of  $\mathbf{U}$  given by

$$\mathbf{U}\mathbf{P} \approx \mathbf{Q} [\mathbf{R}_{11} \mid \mathbf{R}_{12}], \quad (\text{A.1})$$

where  $\mathbf{P}$  is a  $N \times N$  permutation matrix,  $\mathbf{Q}$  an  $m \times r$  matrix with orthonormal columns,  $\mathbf{R}_{11}$  an  $m \times r$  upper triangular matrix, and  $\mathbf{R}_{12}$  an  $r \times (N-r)$  matrix. In practice, the rank  $r$  is unknown

and thus the pivoted Gram-Schmidt process involved in (A.1) is continued until  $\|\mathbf{U} - \mathbf{U}^c \mathbf{C}\| \leq \epsilon$  is achieved for a predefined accuracy  $\epsilon$ . Given (A.1), an  $r \times (N - r)$  matrix  $\mathbf{T}$  is sought for such that

$$\mathbf{R}_{11} \mathbf{T} = \mathbf{R}_{12}. \quad (\text{A.2})$$

When  $\mathbf{R}_{11}$  is ill-conditioned, [23] suggests a solution  $\mathbf{T}$  with minimum  $\|\mathbf{T}\|_F$ . Using (A.2) in (A.1) we arrive at

$$\mathbf{U} \mathbf{P} \approx \mathbf{Q} \mathbf{R}_{11} [\mathbf{I} \mid \mathbf{T}], \quad (\text{A.3})$$

or, equivalently,

$$\mathbf{U} \approx \mathbf{U}^c [\mathbf{I} \mid \mathbf{T}] \mathbf{P}^T = \mathbf{U}^c \mathbf{C}, \quad (\text{A.4})$$

where  $\mathbf{U}^c$  contains the first  $r$  columns of  $\mathbf{U} \mathbf{P}$  and  $\mathbf{C} = [\mathbf{I} \mid \mathbf{T}] \mathbf{P}^T$ .



## Appendix B

### Theoretical Framework for Bi-Fidelity Approximation Error Bound

This section is devoted to proving Theorem 2.2.1 of Chapter 2. Recall that  $\mathbf{H}$  represents the QoI realization matrix for the high-fidelity model, and  $\mathbf{L}$  the corresponding matrix for the low-fidelity model. This analysis assumes that these matrices are related through matrices  $\mathbf{T}$  and  $\mathbf{E}$  according to (2.7). Following the discussions of Section 2.2.4 and, in particular, the bound in (2.8), we seek to identify conditions implying that  $\|\mathbf{T}\|$  is bounded and  $\|\mathbf{E}\|$  is small.

For  $\tau \geq 0$ , define

$$\epsilon(\tau) := \arg \min_{\epsilon} \left\{ \forall \mathbf{x} \in \mathbb{R}^N : \quad \|\mathbf{H}\mathbf{x}\|^2 \leq \tau \|\mathbf{L}\mathbf{x}\|^2 + \epsilon \|\mathbf{x}\|^2 \right\}. \quad (\text{B.1})$$

Note that  $\epsilon(\tau)$  is well-defined, is a non-increasing function of  $\tau$  and, satisfies  $\epsilon(\tau) \leq \|\mathbf{H}\|^2$ . Defining  $\epsilon$  and  $\tau$  this way is equivalent to the definition (2.9) used in Theorem 2.2.1. We present this equivalence here as a lemma.

**Lemma B.0.1.** *For any  $\tau$ , for  $\epsilon(\tau)$  defined as in (B.1), it follows that*

$$\epsilon(\tau) = \lambda_{\max}(\mathbf{H}^T \mathbf{H} - \tau \mathbf{L}^T \mathbf{L}), \quad (\text{B.2})$$

*i.e.,  $\epsilon(\tau)$  is the smallest  $\epsilon$  such that  $\tau \mathbf{L}^T \mathbf{L} + \epsilon \mathbf{I} - \mathbf{H}^T \mathbf{H}$  is a semi-positive-definite (SPD) matrix.*

*Proof.* For a given  $\tau$ , and  $\epsilon(\tau)$  as from (B.1), it follows that for all nontrivial  $\mathbf{x}$ ,

$$\begin{aligned} \|\mathbf{H}\mathbf{x}\|^2 &\leq \tau \|\mathbf{L}\mathbf{x}\|^2 + \epsilon(\tau) \|\mathbf{x}\|^2; \\ \frac{\tau \|\mathbf{L}\mathbf{x}\|^2 + \epsilon(\tau) \|\mathbf{x}\|^2 - \|\mathbf{H}\mathbf{x}\|^2}{\|\mathbf{x}\|^2} &\geq 0. \end{aligned}$$

By the definition in (B.1),  $\epsilon(\tau)$  is the smallest such number making this inequality hold for all  $\mathbf{x}$ , meaning that the inequality is equality for some  $\mathbf{x}$ , implying

$$\begin{aligned}\lambda_{\min}(\tau \mathbf{L}^T \mathbf{L} + \epsilon(\tau) \mathbf{I} - \mathbf{H}^T \mathbf{H}) &= 0; \\ \lambda_{\min}(\tau \mathbf{L}^T \mathbf{L} - \mathbf{H}^T \mathbf{H}) &= -\epsilon(\tau); \\ \lambda_{\max}(\mathbf{H}^T \mathbf{H} - \tau \mathbf{L}^T \mathbf{L}) &= \epsilon(\tau),\end{aligned}$$

completing the proof.  $\square$

We now identify the specific matrices  $\mathbf{T}$  which we consider for our analysis of (2.7).

**Lemma B.0.2.** *Let the SVD of  $\mathbf{L}$  be given by*

$$\mathbf{L} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T.$$

Let  $\mathcal{V}_k$  be the linear subspace spanned by the first  $k$  singular vectors in  $\mathbf{V}$ , i.e., the right singular vectors associated with the  $k$  largest singular values, and let  $\mathbf{V}_k$  be the matrix of those singular vectors. Let

$$\begin{aligned}\mathbf{P}_{\mathcal{V}_k} &= \mathbf{V}_k \mathbf{V}_k^T; \\ \mathbf{T} &:= \mathbf{H} \mathbf{P}_{\mathcal{V}_k} \mathbf{L}^+, \end{aligned} \tag{B.3}$$

where  $\mathbf{P}_{\mathcal{S}}$  denotes the orthogonal projection matrix onto  $\mathcal{S}$ , and  $^+$  denotes the Moore-Penrose pseudo-inverse. Then for  $k \leq \text{rank}(\mathbf{L})$ ,

$$\mathbf{H} - \mathbf{T} \mathbf{L} = \mathbf{H} \mathbf{P}_{\mathcal{V}_k^\perp}. \tag{B.4}$$

*Proof.* Let  $\mathcal{N}(\mathbf{L})$  denote the nullspace of  $\mathbf{L}$ . Note that

$$\mathbf{L}^+ \mathbf{L} = \mathbf{P}_{\mathcal{N}(\mathbf{L})^\perp}.$$

Since  $k \leq \text{rank}(\mathbf{L})$ , then  $\mathcal{V}_k \subset \mathcal{N}(\mathbf{L})^\perp$ . Then from (B.3), it follows that

$$\begin{aligned}
 \mathbf{H} - \mathbf{T}\mathbf{L} &= \mathbf{H} - \mathbf{H}\mathbf{P}_{\mathcal{V}_k}\mathbf{L}^\perp\mathbf{L}, \\
 &= \mathbf{H}(\mathbf{I} - \mathbf{P}_{\mathcal{V}_k}\mathbf{P}_{\mathcal{N}(\mathbf{L})^\perp}) \\
 &= \mathbf{H}(\mathbf{I} - \mathbf{P}_{\mathcal{V}_k}), \\
 &= \mathbf{H}\mathbf{P}_{\mathcal{V}_k^\perp}.
 \end{aligned}$$

□

This result allows us to bound the two key quantities,  $\|\mathbf{E}\| = \|\mathbf{H} - \mathbf{T}\mathbf{L}\|$  and  $\|\mathbf{T}\|$  in terms of  $\epsilon(\tau)$ .

**Lemma B.0.3.** *Let  $\tau$  and  $\epsilon(\tau)$  satisfy (B.1), and let  $\mathbf{T}$  be as defined in (B.3). Then, for  $\sigma_k$ , the  $k$ th largest singular value of  $\mathbf{L}$ ,*

$$\|\mathbf{E}\|^2 = \|\mathbf{H} - \mathbf{T}\mathbf{L}\|^2 \leq \tau\sigma_{k+1}^2 + \epsilon(\tau), \quad (\text{B.5})$$

$$\|\mathbf{T}\|^2 \leq \tau + \epsilon(\tau)\sigma_k^{-2}. \quad (\text{B.6})$$

*Proof.* For any  $\mathbf{x} \in \mathbb{R}^N$  with  $\|\mathbf{x}\| \leq 1$ ,

$$\begin{aligned}
 \|(\mathbf{H} - \mathbf{T}\mathbf{L})\mathbf{x}\|^2 &= \|\mathbf{H}\mathbf{P}_{\mathcal{V}_k^\perp}\mathbf{x}\|^2, \\
 &\leq \tau\|\mathbf{L}\mathbf{P}_{\mathcal{V}_k^\perp}\mathbf{x}\|^2 + \epsilon(\tau)\|\mathbf{P}_{\mathcal{V}_k^\perp}\mathbf{x}\|^2, \\
 &\leq \tau\sigma_{k+1}^2 + \epsilon(\tau),
 \end{aligned}$$

which proves (B.5). We have constructed  $\mathbf{T}$  so that

$$\mathbf{T}\mathbf{y} = \mathbf{0}, \quad \mathbf{y} \in \mathcal{R}(\mathbf{L})^\perp; \quad (\text{B.7})$$

$$\mathbf{T}\mathbf{L}\mathbf{y} = \mathbf{0}, \quad \mathbf{y} \in \mathcal{V}_k^\perp, \quad (\text{B.8})$$

where  $\mathcal{R}(\mathbf{L})$  is the range of  $\mathbf{L}$ . It follows that

$$\begin{aligned}
\|\mathbf{T}\|^2 &= \max_{\|\mathbf{x}\| \leq 1} \|\mathbf{T}\mathbf{x}\|^2, \\
&\stackrel{\text{(B.7)}}{=} \max_{\|\mathbf{L}\mathbf{y}\| \leq 1} \|\mathbf{T}\mathbf{L}\mathbf{y}\|^2, \\
&\stackrel{\text{(B.8)}}{=} \max_{\|\mathbf{L}\mathbf{y}\| \leq 1, \mathbf{y} \in \mathcal{V}_k} \|\mathbf{H}\mathbf{y}\|^2, \\
&\leq \tau \|\mathbf{L}\mathbf{y}\|^2 + \epsilon(\tau) \|\mathbf{y}\|^2,
\end{aligned}$$

for some  $\mathbf{y}$  achieving the maximum, and the last inequality uses the definition of  $\epsilon(\tau)$ . Note that  $\|\mathbf{L}\mathbf{y}\| \leq 1$ , and for  $\mathbf{y} \in \mathcal{V}_k$ ,  $\|\mathbf{y}\| \leq \sigma_k^{-1} \|\mathbf{L}\mathbf{y}\| \leq \sigma_k^{-1}$ . Therefore,

$$\|\mathbf{T}\|^2 \leq \tau + \epsilon \sigma_k^{-2},$$

which proves (B.6). □

The proof of Theorem 2.2.1 follows from the combination of the bound provided in (2.8) and the results of Lemma B.0.3. To explain the minimization of (2.11) over  $k$ , we note that the matrix  $\mathbf{T}$  in (B.0.2) can be constructed for any rank  $k$  not larger than the rank of the low-fidelity data matrix  $\mathbf{L}$ .

As a brief remark we note that (2.11) may instead be taken as

$$\|\mathbf{H} - \hat{\mathbf{H}}\| \leq \min_{\tau_1, \tau_2, k \leq \text{rank}(\mathbf{L})} (1 + \|\mathbf{C}_L\|) \sqrt{\tau_1 \sigma_{k+1}^2 + \epsilon(\tau_1)} + \|\mathbf{L} - \hat{\mathbf{L}}\| \sqrt{\tau_2 + \epsilon(\tau_2) \sigma_k^{-2}}, \quad (\text{B.9})$$

where the choices  $(\tau_1, \epsilon_1(\tau_1))$  and  $(\tau_2, \epsilon_2(\tau_2))$  both satisfy (B.1). Choosing different values may produce more effective values depending on  $\sigma_k$ ,  $\sigma_{k+1}$ ,  $\|\mathbf{C}_L\|$ , and  $\|\mathbf{L} - \hat{\mathbf{L}}\|$ . For simplicity, we have restricted the results of this work to the optimization over a single point  $(\tau, \epsilon(\tau))$ .

## Appendix C

### Computational Methods for Computing Eigenvalue Decompositions

#### C.1 Computing Eigenvalues and Eigenvectors

In this section we provide three algorithms of interest for forming the eigenvalue decomposition of the matrix  $\tilde{\mathbf{H}}$ , as in (5.9): Lanczos method, a randomized version of Lanczos method, and randomized SVD.

##### C.1.1 Lanczos Algorithm

Lanczos is an iterative method that forms a basis  $\mathbf{Q}$  for the Krylov space  $\{\mathbf{x}, \tilde{\mathbf{H}}\mathbf{x}, \dots, \tilde{\mathbf{H}}^r\mathbf{x}\}$  to obtain the decomposition

$$\tilde{\mathbf{H}} \approx \mathbf{Q}_r \mathbf{T}_r \mathbf{Q}_r^T$$

where  $\mathbf{Q}_r$  is an  $N \times r$  matrix with orthonormal columns and  $\mathbf{T}$  is an  $r \times r$  tridiagonal matrix whose eigenpairs  $(\lambda_i, \mathbf{s}_i)$  are easily calculated. Then the eigenpairs of  $\tilde{\mathbf{H}}$  are  $(\lambda_i, \mathbf{v}_i)$  where  $\mathbf{v}_i = \mathbf{Q}_r \mathbf{s}_i$ .

Algorithm 4, based on work in [6], provides the pseudocode for the Lanczos method with a reorthogonalization step for improved stability. Note, we would like to highlight that the most expensive lines of the algorithm are the matrix-vector product involving  $\tilde{\mathbf{H}}$ . For large  $N$ ,  $\tilde{\mathbf{H}}$  is not formed, rather these steps involve either matrix-vector products with each component of  $\tilde{\mathbf{H}}$ , or matrix-free calculations involving components of  $\tilde{\mathbf{H}}$ .

**Algorithm 4:** Lanczos Algorithm for Identifying Eigenpairs

**Given** matrix  $\tilde{H} \in \mathbb{R}^{N \times N}$ , random vector  $\mathbf{x} \in \mathbb{R}^N$ , and tolerance  $TOL$

**Returns** eigenpairs  $\mathbf{\Lambda}, \mathbf{V}$

$\mathbf{q} = \mathbf{x} / \|\mathbf{x}\|$

$\mathbf{Q}_1 = \mathbf{q}$

$\mathbf{r} = \tilde{H}\mathbf{q}$

$a_1 = \mathbf{q}^T \mathbf{r}$

$\mathbf{r} = \mathbf{r} - a_1 \mathbf{q}$

$b_1 = \|\mathbf{r}\|$

$i = 1$

**while**  $b_i > TOL$  **do**

$i = i + 1$

$\mathbf{v} = \mathbf{q}$

$\mathbf{q} = \mathbf{r} / b_{i-1}$

$\mathbf{Q}_i = \mathbf{q}$

$\mathbf{r} = \tilde{H}\mathbf{q} - b_{i-1}\mathbf{v}$

$a_i = \mathbf{q}^T \mathbf{r}$

$\mathbf{r} = \mathbf{r} - a_i \mathbf{q}$

    (reorthogonalization step:)

$\mathbf{r} = \mathbf{r} - \mathbf{Q}(\mathbf{Q}^T \mathbf{r})$

$b_i = \|\mathbf{r}\|$

**end**

$\mathbf{T} = \text{tridiag}(\mathbf{a}, \mathbf{b}, \mathbf{b})$

$[\mathbf{S}, \mathbf{\Lambda}] = \text{eig}(\mathbf{T})$

$\mathbf{V} = \mathbf{Q}\mathbf{S}$

### C.1.2 Randomized Version of Lanczos Algorithm

In a randomized version of Lanczos algorithm (see [57, 18]), Lanczos algorithm is performed on a reduced problem. Algorithm 5 provides the pseudocode for this method. The main difference between this method and the non-randomized version is the initial phase and final phase.

In the initial phase we form an  $N \times \ell$  Gaussian matrix  $\mathbf{\Omega}$ , and identify orthogonal basis  $\hat{\mathbf{Q}}$  from  $\mathbf{Y} = \tilde{\mathbf{H}}\mathbf{\Omega}$ . Typically  $\ell = 2r$  to provide oversampling. Lanczos iteration is then performed on  $\mathbf{B} = \hat{\mathbf{Q}}^T \mathbf{Y} (\hat{\mathbf{Q}}^T \mathbf{\Omega})^{-1} \approx \hat{\mathbf{Q}}^T \tilde{\mathbf{H}} \hat{\mathbf{Q}}$  to obtain decomposition

$$\mathbf{B} = \mathbf{Z} \mathbf{\Lambda} \mathbf{Z}^T.$$

Then it follows that

$$\tilde{\mathbf{H}} \approx \hat{\mathbf{Q}} \mathbf{Z} \mathbf{\Lambda} \mathbf{Z}^T \hat{\mathbf{Q}}^T.$$

**Algorithm 5:** Randomized Lanczos Algorithm for Identifying Eigenpairs

**Given** matrix  $\tilde{H} \in \mathbb{R}^{N \times N}$ , rank  $\ell$ , and tolerance  $TOL$

**Returns** eigenpairs  $\Lambda, V$

$\Omega = \text{randn}(N, \ell)$

$Y = \tilde{H}\Omega$

$[\hat{Q}, \hat{R}] = qr(Y)$

$B = \hat{Q}^T Y (\hat{Q}^T \Omega)^{-1}$

$x = \text{randn}(\ell, 1)$   $q = x / \|x\|$

$Q_1 = q$

$r = Bq$

$a_1 = q^T r$

$r = r - a_1 q$

$b_1 = \|r\|$

$i = 1$

**while**  $b_i > TOL$  **do**

$i = i + 1$

$v = q$

$q = r / b_{i-1}$

$Q_i = q$

$r = Bq - b_{i-1} v$

$a_i = q^T r$

$r = r - a_i q$

    (reorthogonalization step:)

$r = r - Q(Q^T r)$

$b_i = \|r\|$

**end**

$T = \text{tridiag}(a, b, b)$

$[S, \Lambda] = \text{eig}(T)$

$V = \hat{Q}QS$



## C.2 Comparison of Computational Cost

This section discusses the computational cost of four different approaches of this work, based on a high-fidelity mapping which uses a fine grid solver with  $N$  deterministic degrees of freedom, and a bi-fidelity mapping, which relies on both a fine grid solver with  $N$  deterministic degrees of freedom, and a coarse grid solver with  $n$  deterministic degrees of freedom.

Table C.1 provides a floating-point operation (FLOP) comparison for the four methods for forming a rank  $r$  approximation to the posterior covariance: (i) randomized Lanczos, (ii) randomized SVD, (iii) bi-fidelity approach with randomized Lanczos, and (iv) bi-fidelity approach with randomized SVD. The steps include the FLOPs to form components of the rank  $r$  eigenvalue decomposition, e.g.,  $\tilde{\mathbf{H}} \approx \mathbf{V}_r \mathbf{\Lambda}_r \mathbf{V}_r^T$ , as well as the matrix  $\mathbf{D}_r$  when nontrivial, i.e., for the bi-fidelity approximation.

Let  $\mathcal{C}_F(N)$  and  $\mathcal{C}_A(N)$  denote a forward and adjoint solve, respectively, with  $N$  deterministic degrees of freedom for the solver. We leave it in this form, as this cost will depend on type of solver utilized. We also note that  $\mathbf{\Omega}$  is an  $N \times r$  Gaussian matrix. Total approximations are provided with respect to the high-fidelity cost, i.e., assuming  $N \gg n, m, r$ .

Method	Steps	FLOPs
Randomized	$\mathbf{Y} = \tilde{\mathbf{H}}\mathbf{\Omega}$	$r\mathcal{C}_F(N), r\mathcal{C}_A(N)$
Lanczos	$QR(\mathbf{Y}) = \mathbf{Q}\mathbf{R}$	$2r^2N - 2r^3/3$
	$\mathbf{W} = \mathbf{Q}^T\mathbf{Y}(\mathbf{Q}^T\mathbf{\Omega})^{-1}$	$4r^2N + 3r^3 - 2r^2 + r$
	$\mathbf{W} = \mathbf{U}_r\mathbf{\Lambda}_r\mathbf{U}_r^T$ , via Lanczos of $r \times r$	$2r^3 - r^2$
	$\mathbf{V}_r := \mathbf{Q}\mathbf{U}_r$	$2r^2N - rN$
Total:		$\approx N(8r^2 - 1) + r\mathcal{C}_F(N) + r\mathcal{C}_A(N)$
Randomized	$\mathbf{Y} = \mathbf{\Gamma}_{\text{noise}}^{-1/2}\mathbf{A}\mathbf{\Gamma}_{\text{prior}}^{1/2}\mathbf{\Omega}$	$r\mathcal{C}_F(N)$
SVD	$QR(\mathbf{Y}) = \mathbf{Q}\mathbf{R}$	$2r^2m - 2r^3/3$
	$\mathbf{Z} = \mathbf{Q}^T\mathbf{\Gamma}_{\text{noise}}^{-1/2}\mathbf{A}\mathbf{\Gamma}_{\text{prior}}^{1/2}$	$r\mathcal{C}_A(N)$
	$\mathbf{Z} = \mathbf{U}_r\mathbf{\Lambda}_r\mathbf{V}_r^T$ , via SVD of $r \times N$	$2r^2N + 2r^3$
Total:		$\approx 2r^2N + r\mathcal{C}_F(N) + r\mathcal{C}_A(N)$
Bi-Fidelity	$\mathbf{\Gamma}_{\text{noise}}^{-1/2}\mathbf{B}\mathbf{\Gamma}_B\mathbf{B}^T\mathbf{\Gamma}_{\text{noise}}^{-1/2} \approx \mathbf{U}_r\mathbf{S}_r\mathbf{U}_r^T$	$\approx n(8r^2 - 1) + r\mathcal{C}_F(n) + r\mathcal{C}_A(n)$
Approach	$\hat{\mathbf{V}}_r = \mathbf{\Gamma}_{\text{prior}}^{1/2}\mathbf{A}^T\mathbf{\Gamma}_{\text{noise}}^{-1/2}\mathbf{U}_r(\mathbf{S}_r^+)^{1/2}$	$rm + r\mathcal{C}_A(N)$
(Lanczos)	$\hat{\mathbf{D}}_r = (\mathbf{I} + \mathbf{S}_r\hat{\mathbf{V}}_r^T\hat{\mathbf{V}}_r)^{-1}\mathbf{S}_r$	$r(r+1)N + r^3 + 3r^2/2 + 3r/2$
Total:		$\approx r^2N + rN + r\mathcal{C}_A(N)$
Bi-Fidelity	$\mathbf{\Gamma}_{\text{noise}}^{-1/2}\mathbf{B}\mathbf{\Gamma}_B^{1/2} \approx \mathbf{U}_r\mathbf{\Sigma}_r\mathbf{W}_r^T$	$\approx 2r^2n + r\mathcal{C}_F(n) + r\mathcal{C}_A(n)$
Approach	$\hat{\mathbf{V}}_r = \mathbf{\Gamma}_{\text{prior}}^{1/2}\mathbf{A}^T\mathbf{\Gamma}_{\text{noise}}^{-1/2}\mathbf{U}_r(\mathbf{S}_r^+)^{1/2}$	$rm + r\mathcal{C}_A(N)$
(SVD)	$\hat{\mathbf{D}}_r = (\mathbf{I} + \mathbf{S}_r\hat{\mathbf{V}}_r^T\hat{\mathbf{V}}_r)^{-1}\mathbf{S}_r$	$r(r+1)N + r^3 + 3r^2/2 + 3r/2$
Total:		$\approx r^2N + rN + r\mathcal{C}_A(N)$

Table C.1: Number of FLOPs for Different Low-Rank Approaches