Preface

Parameter studies are everywhere in computational science. Complex engineering simulations must run several times with different inputs to effectively study the relationships between inputs and outputs. Studies like optimization, uncertainty quantification, and sensitivity analysis produce sophisticated characterizations of the input/output map. But thorough parameter studies are more difficult when each simulation is expensive and the number of parameters is large. In practice, the engineer may try to limit a study to the most important parameters, which effectively reduces the *dimension* of the parameter study.

Active subspaces offer a more general approach to reduce the study's dimension. They identify a set of important directions in the input space. If the engineer discovers a model's low-dimensional active subspace, she can exploit the reduced dimension to enable otherwise infeasible parameter studies for expensive simulations with many input parameters. This book develops active subspaces for dimension reduction in parameter studies.

My journey with active subspaces started at Stanford's Center for Turbulence Research Summer Program in 2010 with Gianluca Iaccarino (Stanford), Alireza Doostan (CU Boulder), and Qiqi Wang (MIT). We were studying surrogate models for calibrating the six inputs of an expensive scramjet simulation with Markov chain Monte Carlo. After struggling for two weeks to get the chain to converge, Qiqi came to the morning meeting touting some "magic parameters." He had found three orthogonal directions in the six-dimensional parameter space such that small input perturbations along these directions barely changed the pressure at prescribed sensor locations. The remaining three orthogonal directions gave three linear combinations of the six parameters; perturbing the latter linear combinations significantly changed the pressure, so they were more important for the calibration. He had found these directions by studying the pressure's gradient with respect to the input parameters. Lacking a technically descriptive name, Qiqi called the three important linear combinations *magic*. Given the scramjet runs we had available, he could build a regression surface in three variables with a higher degree polynomial than in six variables. We all thought this was a really cool idea.

The work was published in the Summer Program's proceedings and later as an AIAA conference paper [35]. But then it sat on the back burner for a while. In late 2011, I started bugging Qiqi that we should write a journal-worthy version of this *input dimension reduction* idea with more precise statements and formal development. Since so many uncertainty quantification techniques suffered from the dreaded curse of dimensionality, I thought reducing the dimension of a model's inputs could have a big impact. It took a while to get the formulation right, and we went back and forth with the reviewers. Amid that exchange, I found Trent Russi's 2010 Ph.D. thesis [108], in which he proposed a similar idea for reducing a model's input dimension with subspaces derived from random samples of the gradient; he called these subspaces *active subspaces*. We thought that was a catchy and reasonably descriptive name, so we adopted it.

The paper was finally accepted and appeared as [28]. Around the same time, Qiqi and I connected with Youssef Marzouk (MIT) and Tan Bui (UT Austin), who were working on techniques for improving Markov chain Monte Carlo for Bayesian inverse problems—including some subspace-based dimension reduction ideas. We put together a proposal for the DOE's Advanced Scientific Computing Research Applied Mathematics program, and we were awarded a three-year grant to develop active subspaces for data-intensive inverse problems. That grant has supported my effort to write this monograph.

Audience. Computational science is inherently interdisciplinary. I hope that this text offers something useful and interesting to researchers across fields. Dimension reduction must be practical and well-motivated—with easy-to-implement algorithms and easy- to-interpret results—to impact real applications. But the techniques must also be theoretically well-founded with rigorous performance guarantees for cases that satisfy simplifying assumptions. When simplified cases lack the challenges of real applications, engineers accuse the analysts of working on *toy problems*. When an engineering method has no rigorous performance guarantees, the analysts deride the method as *heuristic*—to which the engineer may respond that it works well in practice.

These squabbles are unlikely to subside in the near future, and this book will not resolve them. However, I have tried to balance both objectives. I have tried to develop the material so that it is amenable to analysis. And I have tried to demonstrate the value of active subspaces in engineering applications. I hope that both groups will benefit. There is still plenty to do on both fronts—analysis and engineering. Computational science graduate students may find a new perspective on their research with these techniques and subsequently advance the state of the art.

Outline. The first chapter provides a quick start with active subspaces for engineers paralyzed by their high-dimensional parameter studies. It offers some easy-to-implement procedures for discovering whether the given model admits an active subspace. If these tests show evidence of an active subspace, then the practitioner should be well-motivated to read on. The first chapter is reasonably self-contained. As such, I repeat some of the material in later chapters for further development.

The second chapter puts active subspaces into the broader context of algorithm research in uncertainty quantification. It includes important references to related statistics research on *sufficient dimension reduction* that blossomed more than 20 years ago.

The third and fourth chapters develop the technical content, including defining the active subspace, proposing and analyzing a method to discover the active subspace, and discussing strategies to exploit the active subspace for high-dimensional parameter studies.

The fifth chapter demonstrates the utility of active subspaces in three engineering applications: (i) studying the safe operating regime of a hypersonic scramjet, (ii) characterizing the relationship between model inputs and maximum power in a photovoltaic solar cell model, and (iii) optimizing the shape of an airfoil. Each section has just enough detail to put the application in the mathematical framework for active subspaces.

The quickest way to impact applications is to provide easy-to-use software. I have not included any language-specific implementations in the text—mostly because of the rapid pace of innovation in software and computing. However, I maintain the website `activesubspaces.org`, and I provide scripts and utilities there for working with active subspaces.

Acknowledgments. Throughout the text, I use the first person plural *we*[1] to reflect that the work in this book has been a collaborative effort. Although the choice and arrangement of words and symbols are my own, the intellectual content contains

[1] In the first chapter, I use the second person singular *you* to make directions easier to read.

significant contributions from my esteemed colleagues. My own interests cannot possibly cover the breadth of expertise needed for interdisciplinary computational science research—especially the domain expertise in the applications in Chapter 5. The final work would not have been possible without my colleagues' efforts.

I am very grateful for Qiqi Wang and Eric Dow at MIT. Much of the technical development in Chapters 3 and 4 comes from our paper [28]. David Gleich (Purdue) and I worked together to analyze the random sampling method for estimating active subspaces presented in Chapter 3, and his help with the presentation and several tricky turns in the proofs was invaluable [30]. Youssef Marzouk (MIT) and Tan Bui (UT Austin) helped refine my ideas on Bayesian inverse problems in Chapter 4. The scramjet application from Chapter 5 was a primary objective at Stanford's NNSA-funded Predictive Science Academic Alliance Program center. The specific work with active subspaces was performed jointly with Michael Emory (Stanford), Johan Larsson (UMD College Park), and Gianluca Iaccarino (Stanford) [29]. Their comments from the engineering side helped ensure that active subspaces could be valuable in real applications. The photovoltaics application in Chapter 5 consists of joint work with Brian Zaharatos (Colorado School of Mines) and Mark Campanelli (NREL) [37]. The airfoil shape optimization in Chapter 5 consists of joint work with Trent Lukaczyk, Francisco Palacios, and Juan Alonso at Stanford's Aerospace Design Laboratory [87]. I am especially thankful for Trent's help in running the simulations, preparing the figures, and interpreting the results in section 5.3. Lastly, I am grateful for the helpful suggestions and support from Ralph Smith (NCSU) and the *SIAM Spotlights* anonymous reviewers. Their comments helped me to clarify the presentation.

I have been extremely fortunate to receive funding from the U.S. Department of Energy. As a postdoc at Stanford, I was funded by the National Nuclear Security Administration under Award NA28614 through the Predictive Science Academic Alliance Program. As an assistant professor at Colorado School of Mines, my efforts have been supported by the Office of Science Advanced Scientific Computing Research Applied Mathematics program under Award DE-SC-0011077.