Title
Quantifying the agreement between computational models and experimental data under uncertainty

Author
Hegde, Arun Shantaram

Publication Date
2020

Peer reviewed|Thesis/dissertation

Quantifying the agreement between computational models and experimental data under uncertainty

by

Arun Shantaram Hegde

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Mechanical Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Michael Frenklach, Chair
Professor Kameshwar Poolla
Professor Murat Arcak

Spring 2020

Quantifying the agreement between computational models and experimental data under uncertainty

Abstract

Quantifying the agreement between computational models and experimental data under uncertainty

by

Arun Shantaram Hegde

Doctor of Philosophy in Mechanical Engineering

University of California, Berkeley

Professor Michael Frenklach, Chair

Bound-to-bound data collaboration (abbreviated B2BDC) is a deterministic optimization-based approach for uncertainty quantification. The framework combines models and data from multiple sources by formulating inequality constraints over a parameter space. This dissertation explores the following question: how can agreement between computational models and experimental data be quantified while necessarily accounting for uncertainty in both model parameters and observations? In a typical B2BDC application, this is performed by constructing a *dataset* – a collection of constraints over an uncertain parameter space involving surrogate models, experimental data, and prior knowledge – and then assessing its *consistency*. Our first contribution is a formalization of this procedure within an iterative context. This new strategy effectively extends the applicability of the B2BDC technique and can be viewed as a natural extension of previous work. Oftentimes, demonstrating model-data disagreement is just as important as verifying agreement. In B2BDC, this is manifested through dataset *inconsistency*. Our second contribution is a new tool for analyzing inconsistency called the vector consistency measure. This measure provides a more thorough diagnosis of an inconsistent dataset by computing minimal constraint corrections that lead to consistency. The inclusion of weights facilitates domain expert knowledge and opinions to be incorporated in the process of resolving an inconsistency. The primary developments in this thesis are methodological. Their application is illustrated on various examples, ranging from the small-scale instances drawn from the literature to larger-scale realistic gas combustion datasets.

To my parents and my teachers.

Contents

Contents ii

List of Figures iv

List of Tables vi

1 Introduction 1
- 1.1 Predictive modeling . . . 1
- 1.2 Comparisons of model and data: literature review . . . 1
- 1.3 Structure and contributions . . . 2

2 Review of Bound-To-Bound Data Collaboration 4
- 2.1 Introduction . . . 4
- 2.2 Problem setup and notation . . . 4
- 2.3 Datasets and consistency . . . 5
- 2.4 Scalar consistency measure . . . 6
- 2.5 Prediction and inference . . . 7
- 2.6 Solution mapping and the surrogate dataset . . . 7
- 2.7 Computational strategy for NQCQPs . . . 9
- 2.8 Implications on prediction and inference . . . 14
- 2.9 A two parameter example . . . 15
- 2.10 Chapter summary . . . 18

3 Iterative Construction of a B2BDC Dataset 19
- 3.1 Introduction . . . 19
- 3.2 Bayesian UQ and history matching . . . 20
- 3.3 B2BDC and Bayesian history matching . . . 23
- 3.4 Previous iterative strategies in B2BDC . . . 24
- 3.5 Incorporating waves in B2BDC . . . 24
- 3.6 Example: Calibration of an Aerodynamic Body . . . 25
- 3.7 Discussion and future work . . . 34

4 Iterative Construction of a B2BDC Dataset: Rational Quadratic Surrogate Models **36**
4.1 Introduction . . . 36
4.2 Fitting over $\mathcal{H}$ . . . 36
4.3 Fitting over $\mathcal{F}$ . . . 38
4.4 Sampling RQ datasets . . . 40
4.5 Example continued: calibration of an aerodynamic body . . . 41
4.6 Chapter summary . . . 44

5 Additional Feasibility Criteria **46**
5.1 Introduction . . . 46
5.2 Bounding ellipsoids to feasible samples . . . 47
5.3 Minimum volume enclosing ellipsoids . . . 47
5.4 Generating additional quadratic constraints using support vector machines . . . 50
5.5 Example: Chemical Spill Model . . . 52
5.6 Chapter summary . . . 58

6 Sensitivities and the Scalar Consistency Measure **59**
6.1 Introduction . . . 59
6.2 Running examples . . . 60
6.3 Sensitivity through perturbation . . . 60
6.4 A strategy for resolving inconsistency . . . 62
6.5 Strengthening the relaxation via redundant constraints . . . 63
6.6 Resolving inconsistency in GRI-Mech 3.0 . . . 65
6.7 Resolving inconsistency in DLR-SynG . . . 66
6.8 Chapter summary . . . 69

7 The Vector Consistency Measure **70**
7.1 Introduction . . . 70
7.2 The Vector Consistency Measure . . . 70
7.3 Linear examples . . . 73
7.4 Weighted vector consistency . . . 79
7.5 Using the VCM . . . 80
7.6 Examples: GRI-Mech 3.0 and DLR-SynG . . . 81
7.7 Comparison with the SCM . . . 86
7.8 An extension to rational quadratic surrogates . . . 90
7.9 Consistency measures that count: a possible future direction . . . 92
7.10 Chapter summary . . . 94

8 Conclusion **96**

Bibliography **97**

List of Figures

2.1 QOIs and their associated uncertainty bounds, shown in red. 15
2.2 Surrogate models with the associated training points shown in black. 16
2.3 The feasible set of the dataset with coordinate axes displaying the full prior range. 17
2.4 Prediction of the surrogate model over the feasible set (in red) and over the prior $\mathcal{H}$ (in black). 17

3.1 Measured input signals. 27
3.2 Measured output signals. The intervals show the $N = 65$ QOIs and the corresponding uncertainty bounds. 27
3.3 Left, EEFs after the first wave. Right, EEFs after the eighth wave. 29
3.4 Model outputs for approximately 1000 (ignoring failed evaluations) randomly selected feasible parameter vectors after the kth wave. Only the prior and samples after waves $k = 1, 3, 5 - 8$ are shown, with lighter colors corresponding to later waves. Note that wave 8, shown in off-white, is essentially identical to waves 6 and 7. The measured output signals along with the QOI bounds are shown in red. 30
3.5 Model outputs for 1000 randomly selected feasible parameter vectors after the kth wave. The vertical axis has been extended to accommodate the entire output variation across the feasible parameter samples. 31
3.6 The measured output signals are shown in red. 22 feasible samples are shown in blue. The nonlinear least squares optimal point is shown in green. 32
3.7 Pairwise scatter plot associated with the final batch of feasible samples shown in Figure 3.4. The parameter ordering listed on the vertical axis (top-down) corresponds to the ordering listed on the horizontal axis (left-right). Inner (**I**) and outer (**O**) posterior bounds are shown along the horizontal axis. Note, the axes limits have been standardized to the posterior outer bounds. 34

4.1 Fitting error histograms for each of the surrogate types. 41
4.2 EEF criteria for the three types of surrogate models: quadratic, **RQ-H**, and **RQ-F**. 42
4.3 Model outputs after the kth wave. The model evaluations shown in gray are feasible with respect to the new dataset, which includes QOI #7 fit with an RQ surrogate. 43

4.4 The measured output signals are shown in red. Evaluations of 22 feasible samples are shown in blue. The nonlinear least squares optimal point evaluation is shown in green. 44

5.1 Nominal measurements of the QOIs of the ground truth system $C(s,t;x_\star)$ are displayed in red. 53
5.2 Left, EEFs for the first wave. Right, EEFs for the second wave. 55
5.3 Left, EEFs for the third wave. Right, EEFs for the fourth wave. 55
5.4 QOI traces of 1000 randomly selected feasible parameter vectors after the kth wave (i.e., samples from from $\mathcal{F}^{k+1}$). These traces were generated by directly evaluating Equation (5.11). The experimental bounds are shown in red. 56
5.5 Pairwise scatter plot of 5000 randomly selected feasible parameter vectors. Note that the axes have been set to the corresponding posterior outer bounds. 58

6.1 The SCM sensitivity ranking for GRI-Mech 3.0. 66
6.2 The SCM sensitivity ranking for DLR-SynG. 67
6.3 The SCM sensitivity ranking for DLR-SynG after the first constraint removal. . 68

7.1 Illustration of the counter example. The dashed lines indicate the 1-norm ball with radius $c^\star$. 74
7.2 Histogram of ratios for 10,000 random trials with $n_e = 1$. 76
7.3 Histogram of ratios for 10,000 random trials with $n_e = 4$. 77
7.4 Counter example with weighted linear VCM. 78
7.5 VCM optimal relaxations for three standard weight configurations. 82
7.6 VCM relaxations for QOIs #36 and #36 (blue dots) due to random selection of weights. Other feasible relaxations lie in the shaded gray region. The red region is certified infeasible by the SDP results. 83
7.7 VCM relaxations for the DLR-SynG dataset with unit weights on the QOI bounds and null weights on the parameter bounds. 84
7.8 VCM relaxations for the DLR-SynG dataset with bound weights on the QOIs and null weights on the parameters. 85
7.9 VCM relaxations for the DLR-SynG dataset with interval weights on the QOIs and null weights on the parameters. 85
7.10 VCM relaxations for the DLR-SynG dataset with null weights on both QOI #141 as well as the parameters. The remaining QOIs were assigned unit weights. . . 86
7.11 Count VCM relaxations. 94

List of Tables

3.1 The uncertain parameters and the corresponding prior bounds. 26
3.2 Composition of the dataset after the kth wave. The volume fraction is calculated using a collection of 5×10^6 Latin hypercube samples over the prior region $\mathcal{H}$. Note that in each case $\mathcal{F}^{k+1}$ is nonempty. 29

5.1 Composition of the dataset after each wave. The number of ellipsoids corresponds to the number of GP surrogates that passed the error criteria during the kth wave. The volume fraction is calculated using a baseline collection of 5×10^6 LH samples from $\mathcal{H}$. 56
5.2 Prior bounds and posterior outer bounds on the uncertain parameters at the conclusion of wave 4. 57

7.1 Example weights for a generic constraint: $L - w_L\delta_L \leq f(x) \leq U + w_U\delta_U$ 80

Acknowledgments

First and foremost, I'd like to acknowledge my advisors, Professor Andrew Packard and Professor Michael Frenklach. The foundation of the material discussed in this dissertation was built on their collaboration. I am very grateful to have been one of Andy's students. I entered the program in Fall 2014, around the time when Andy was diagnosed with cancer. Despite all of the medical challenges, we continued to meet as a group practically every week until his passing in September 2019. We miss him deeply, but I am immensely thankful for all of the time, the memories, and the lessons (this, of course, also includes log rolling). Andy always had a very honest approach to research. Whenever we discussed our work, be it during our normal group meetings or when traveling, he would emphasize the importance of clear thinking and simple examples. Andy's tireless dedication and tenacity are among the many things that I will carry with me. From Michael, I learned the excitement of science. Michael's advice on writing, presentation, and the importance of telling an overarching story is something that has stuck with me. Similarly, I have found much value in the many kernels of wisdom he has shared over the years – for example, his advice to relax and watch a funny movie the night before a big day. To be succinct, I am really fortunate to have had both Andy and Michael as guides. And as with most things, this has become evermore clear in hindsight. I also very much appreciate the support, consideration, and comments from my dissertation committee members, Professor Kameshwar Poolla and Professor Murat Arcak.

Over the years, I have had the good fortune of learning from many excellent professors and teachers. I am deeply indebted to my undergraduate advisor, Professor Jiong Tang at the University of Connecticut, for his guidance and encouraging me to pursue a PhD. I would also like to thank Professor Philip Smith at the University of Utah for providing us with numerous opportunities to present and share our work; Dr. Sean Smith for the many discussions via webcam and during our visits; and Dr. Bart van Bloemen Waanders for his mentorship and career advice during an internship at Sandia National Laboratories.

In my immediate research group, I would like to thank Jim Oreluk for the innumerable discussions on UQ and practically every other topic; Wenyu Li for the varied chats ranging from research to movies to sports; and Zhenyuan Liu for the similarly varied exchanges on books, podcasts, and soccer. I'd also like to thank Alex Frank for all of the wisdom and support. In the BCCI group and beyond, I'd like to thank Chris Meissen, Jared Porter, Emily Jensen, Octavio Narváez-Aroche, Emmanuel Sin, Kate Schweidel, He (Galaxy) Yin, and Akhil Shetty for all of the friendship, camaraderie, and conversations. To Emmanuel, Kate, Galaxy, and Jared: I am truly grateful for your friendship during this final year (and the previous years too, let's not forget those). I'd also like to thank Johanna Sedman for all her kindness and being so welcoming. Andy always said that one of the most incredible and important things about graduate school is who you meet. The people above are a testament to this.

Finally, I'd like to thank my family. My parents and my brother have always supplied a very patient and unending stream of support and encouragement. They have been with me every step of the way and that single fact has made all the difference.

The material in this dissertation was based on work supported by the U.S. Department of Energy, National Nuclear Security Administration, under award DE-NA0002375. I am deeply grateful to have had this opportunity.

Chapter 1

Introduction

1.1 Predictive modeling

Mathematical modeling lies at the heart of scientific inquiry and engineering decision-making. A general objective of science is to construct models of real-world systems that: (i) agree with available experimental evidence; (ii) provide insight into the inner workings of the system; and (iii) reliably predict unobserved phenomena or future experiments. In many cases, these models are constructed from either first principle or phenomenological laws, and often lead to forms parameterized by partially unknown but meaningful quantities.

Addressing the above points (i), (ii), and (iii) requires a plausible accounting of uncertainty. It is a preeminent fact that uncertainty, or lack of knowledge, pervades all aspects of the scientific process. For example, Saltelli and coauthors emphasize in [101] that: "uncertainty is not an accident of the scientific method, but its substance." In recent years, the topics of validation and verification at the intersection of physical modeling and data analysis have undergone much scrutiny [27, 84, 7]. Establishing the level of agreement between a computational model and experimental data, while transparently managing the uncertainties contributed by both aspects, becomes paramount – a necessary step before prediction can even be justified. As such, these topics of validation, verification, and uncertainty quantification have become a fundamental activity in modern science and engineering.

1.2 Comparisons of model and data: literature review

In settings where uncertainties in both the model and data abound, assessing the match between a computational model and experimental data can become a significant challenge. In practice, this comparison is usually accomplished by examining numerical differences between quantities of interest (QOIs) selected from model outputs and measured data. How uncertainty is framed mathematically, however, can lead to different approaches.

Several common strategies for model validation and prediction are probabilistic and employ a Bayesian inferential framework. These approaches are often built on the foundational

work of Kennedy and O'Hagan [63]. Applications and extensions of this can be found, for example, in studies by Higdon et al. [56, 55] and Bayarri et al. [9, 10]. A common feature among many of these methods is the usage of flexible and easy-to-evaluate stochastic emulators as surrogates for expensive computer models, e.g., as promoted by Sacks et al. [100]. Computational challenges with these approaches have inspired new methods. For example, *Bayesian history matching* [28, 29, 116, 117] aims for tractability by dropping the requirement of a completely probabilistic specification and combines emulation with iterations. Rather than characterizing an entire posterior distribution, the technique seeks regions of the parameter space for which model outputs adequately agree with experimental data.

In certain scenarios, non-traditional or even less nuanced descriptions of uncertainty may be preferable [78]. One such approach, where uncertainty is modeled by set membership constraints, is present in a number of fields, including control theory [33, 108, 65, 14, 8], system identification [57, 119, 25, 81, 77, 62], robust optimization [12], and computational biology [87, 93, 89]. The methodology of *bound-to-bound data collaboration* (B2BDC) [42, 37, 36, 103] is designed to handle questions of model validation and uncertainty quantification in such settings. In contrast with the aforementioned techniques, B2BDC operates by collecting models, experimental bounds, and prior information into what is termed a *dataset* and then formulating constrained optimization problems to quantify model-data agreement and make predictions. Although the resulting expressions tend to be nonconvex with potentially many local optima, the methodology leverages several standard techniques to relax the formulations into efficiently solvable semidefinite programs [19, 47, 90]. This strategy has been effectively applied in several scientific domains, including combustion science [43, 37, 98, 91, 105, 40, 60], atmospheric chemistry [106], and system biology [36, 35, 122].

1.3 Structure and contributions

The underlying challenge that we grapple with in each of the chapters below is how to appropriately quantify model-data agreement under uncertainty. In this dissertation, we take the view of B2BDC, which frames these questions within the context of a *dataset* and its *consistency*. As described in the previous section, this framework examines model validation and uncertainty quantification from a deterministic optimization-based perspective. The basic methodology and computational underpinnings are reviewed in Chapter 2, which serves as a springboard into the remainder of the work. The primary content of this thesis can be divided into two halves:

- The first half consists of Chapters 3-5. These sections focus on the iterative construction of a dataset. In particular, Chapter 3 draws connections between previous B2BDC works and Bayesian history matching, and proposes a formal procedure for iteratively building a B2BDC dataset. Chapter 4 extends this procedure to including rational quadratic surrogate models. Chapter 5 further discusses how additional criteria for measuring model-data agreement can be incorporated through computing bounding

ellipsoids. The key contributions of these sections are the links between B2BDC and history matching, and the resulting iteration-based technique as applied in the B2BDC framework.

- The second half consist of Chapters 6 and 7. These sections focus on *consistency measures*, fundamental tools for quantifying model-data agreement and disagreement. Chapter 6 elaborates on the usage of a previous tool for analyzing dataset consistency developed by Feeley et al. in [35]. By example, we demonstrate the necessity for alternate consistency measures to handle conflicts among models and data. Chapter 7 discusses a new tool, proposed in our recent work [54], that tackles these questions. The key contributions of these sections are a formalization of the previous sensitivity-based strategy as well as a new tool (and framework) for diagnosing and managing dataset inconsistency.

Chapter 2

Review of Bound-To-Bound Data Collaboration

2.1 Introduction

This chapter provides an overview of bound-to-bound data collaboration (B2BDC) and reviews the background information necessary for the remainder of the thesis. Much of the content can be found in earlier works, with the primary reference being the dissertation of Ryan Feeley [35]. As such, this section builds the foundation for new material in subsequent chapters. Emphasis is placed on items relevant to these later sections and departures from previous work are highlighted.

2.2 Problem setup and notation

The framework of B2BDC generally requires the following:

- a pairing between a real-world system $\mathcal{R}$ and a corresponding parameter-dependent mathematical model $\mathcal{M}$;
- the form and structure of $\mathcal{M}$ to be known, with uncertainty primarily stemming from incomplete knowledge of the model parameters;
- availability of experimental data with expert-assessed uncertainties (e.g., as measured from $\mathcal{R}$); and
- numerical instantiations of $\mathcal{M}$ taking as input uncertain parameters and producing outputs corresponding to the experimental observations.

As indicated in the name, B2BDC characterizes the above-described uncertainties through bound constraints. Experimental measurements of N quantities of interest (QOIs) and their uncertainties are encoded by intervals $\{[L_i, U_i]\}_{i=1}^{N}$ rather than individual numerical values.

Prior information representing the current knowledge of the model parameters is denoted by a set $\mathcal{H} \subseteq \mathbb{R}^n$, where n is the dimension of the parameter space. In keeping with notation from earlier works, the variable $x \in \mathbb{R}^n$ is used to denote the parameter vector. The computable mapping described in the final bullet is expressed by a multi-output function $f : \mathbb{R}^n \to \mathbb{R}^N$. Note that f produces simulated versions of the QOIs. Hence f serves as a platform through which the model parameters x can interact with the observed data (experimental bounds), thus providing a basis for comparison between $\mathcal{M}$ and $\mathcal{R}$.

To illustrate f, consider a dynamical system represented by a generic first-order ordinary differential equation:

$$\dot{z}(t) = g(z(t), t, x), \quad z(0) = z_0. \tag{2.1}$$

where $z(t) \in \mathbb{R}^d$ is the state vector and $x \in \mathbb{R}^n$ denotes the collection of model parameters. The N QOIs described above may correspond to certain experimentally-measurable properties of the solution $z(t)$, such as the jth peak response $z_{\text{peak},j} = \max_{t \geq 0} |z_j(t)|$, the jth steady state value $z_{\text{ss},j} = \lim_{t \to \infty} z_j(t)$, etc. Therefore, the parameter-to-QOI map f subsumes the numerical solution of Equation (2.1) and links parameter choices x to simulated versions of the experimental QOIs. A simple illustration of this is provided in Section 2.9.

We note that the experimental bounds, prior information, and models are usually interpreted as "tentatively entertained" in the spirit of Box and Hunter [16]. Moreover, the bounds (both experimental and prior) are generally expert-assessed rather than drawn purely from raw data. Hence expert opinions on items such as measurement reliability, interpretation of experimental records, unmodeled physics, etc. can be factored into the analysis by appropriately inflating or shifting the bounds.

2.3 Datasets and consistency

The central concept in B2BDC is the notion of a *dataset* and its corresponding *feasible set*. The former is a collection of constraints tying together prior information, QOI models, and experimental bounds. The latter is the set of parameters carved out by those constraints. For an individual QOI i, the associated set of feasible parameters is defined by constructing the model-data constraint,

$$\mathcal{F}_i = \{x \in \mathbb{R}^n : x \in \mathcal{H},\ L_i \leq f_i(x) \leq U_i\}. \tag{2.2}$$

The reduction from $\mathcal{H}$ to $\mathcal{F}_i$ is exactly what is learned from the i-th measurement in the context of the model. Accumulating and intersecting all N of such sets produces the complete feasible set,

$$\mathcal{F} = \cap_{i=1}^{N} \mathcal{F}_i, \tag{2.3}$$

which represents the collective information gathered by the dataset. Since $\mathcal{H}$ characterizes the a priori uncertainty in the model parameters, $\mathcal{F}$ is then a set description of the a posteriori uncertainty. The reduction from $\mathcal{H}$ to $\mathcal{F}$ is then precisely due to the inclusion of all observed data.

Any feasible parameter vector $x \in \mathcal{F}$ complies with the prior knowledge and leads to agreement between model predictions and experimental measurements for the designated QOIs. Without additional criteria, feasible parameters are all tentatively valid and cannot be distinguished. If the feasible set is nonempty, i.e., $\mathcal{F} \neq \varnothing$, then the dataset is said to be *consistent*. Conversely, an empty feasible set $\mathcal{F} = \varnothing$ demonstrates that the dataset is *inconsistent* – no satisfactory x can be found. Consistency, or lack thereof, is therefore a property of a dataset and its collective content rather than any individual component.

We note that the nomenclature of "data collaboration" in B2BDC originates in the idea that a dataset is often comprised of diverse, heterogeneous data contributed from separate and independent reference materials, e.g., research laboratories, technical publications, experimental records, etc. The "collaboration" aspect is essentially anonymous. Hence, a dataset represents a community's collective research portfolio and draws from sources discipline-wide in order to corroborate the entire prior-model-data interaction [42, 40]. Consistency analysis can therefore be interpreted as a form of model checking in that a diagnosis of inconsistency confirms that the components are fundamentally in disagreement. Importantly, the nature of this conflict – e.g., whether it stems from a mutual incompatibility among the QOI models, among the data, or both – remains to be ascertained. We may still ask, for instance, does the data invalidate the model or does the model invalidate the data? How to answer such questions very much rests in the hands of domain experts. Developing tools to aid in this line of inquiry and to expose potential sources of inconsistency will be visited in Chapter 6 and Chapter 7. Briefly, we note that a dataset being consistent is not a confirmation of model validity, it only states that the model is capable of reproducing the available data within the reported uncertainty. In general, an unconditional declaration that a model is "valid" is impossible as finite data is only ever available.

2.4 Scalar consistency measure

The consistency of a dataset can be rigorously quantified by computing what is termed a *consistency measure*. The scalar consistency measure (SCM), originally proposed in work by Feeley et al. [37], is one such device:

$$\begin{aligned} C_{\mathrm{D}} := \max_{\gamma,\ x\in\mathcal{H}} \quad & \gamma \\ \text{s.t.} \quad & L_i + \frac{(U_i - L_i)}{2}\gamma \leq f_i(x) \leq U_i - \frac{(U_i - L_i)}{2}\gamma \\ & \text{for } i = 1, \ldots, N. \end{aligned} \tag{2.4}$$

In the above problem, the auxiliary scalar variable γ acts as either a hypothetical tightening or relaxing of all experimental bounds and signals whether the dataset is consistent or not. If $\gamma \geq 0$, then all experimental bounds can be tightened, implying the existence of a feasible $x \in \mathcal{H}$. If $\gamma < 0$, then the experimental bounds must be relaxed (widened) in order to accommodate an $x \in \mathcal{H}$; the dataset is inconsistent. Hence the sign of C_D computed by

Equation (2.4) determines the status of the dataset. Details on the computational strategy used to solve Equation (2.4) are discussed in Section 2.7 and in greater detail in Chapter 6.

2.5 Prediction and inference

A consistent dataset is tentatively valid and the existence of the feasible set $\mathcal{F}$ establishes confidence that the model being studied is suitable for prediction. Let $f_P : \mathbb{R}^n \to \mathbb{R}$ denote a prediction function. For example, f_P may refer to another feature of $\mathcal{M}$ which is of interest to the scientist or decision-maker. As discussed previously, feasible parameters $x \in \mathcal{F}$ cannot be distinguished; they are all candidate solutions. Hence reporting an individual evaluation of $f_P(x)$ for a particular $x \in \mathcal{F}$ is not a justifiable mechanism for prediction. Rather, B2BDC accounts for the indistinguishability by computing the extent of f_P across the entirety of $\mathcal{F}$:

$$\left[\min_{x\in\mathcal{F}} f_P(x),\ \max_{x\in\mathcal{F}} f_P(x)\right]. \tag{2.5}$$

In this way the uncertainty in prediction is characterized by the extreme values, thus accounting for all parameters values judged plausible by the dataset. The overall framework for prediction is referenced in the "bound-to-bound" nomenclature of B2BDC – bounds on the parameters and bounds on the observations are transformed into bounds on the predicted quantity by way of the feasible set.

Different choices of f_P can lead to different inferences. For example, posterior bounds on the parameters can be formulated by projecting $\mathcal{F}$ into the coordinate axes,

$$\left[\min_{x\in\mathcal{F}} x_j,\ \max_{x\in\mathcal{F}} x_j\right] \quad \text{for } j = 1, ..., n. \tag{2.6}$$

The resulting hyperrectangle may then be compared to $\mathcal{H}$. Another useful computation is that of "blind" prediction,

$$\left[\min_{x\in\mathcal{F}_{-i}} M_i(x),\ \max_{x\in\mathcal{F}_{-i}} M_i(x)\right]. \tag{2.7}$$

where $\mathcal{F}_{-i}$ refers to the feasible set with the ith model-data constraint removed. Comparing the blind prediction for a particular QOI, say QOI i, to the corresponding experimental uncertainty $[L_i, U_i]$ provides a measure of how influential the associated constraint is to the dataset. If the blind prediction interval is completely contained within the experimental bounds, then that experiment contributes nothing to our current state of knowledge (of x). By this, we mean that the information in the model-data constraint $L_i \le M_i(x) \le U_i$ is already implied by $\mathcal{F}_{-i}$, i.e., other experiments in the dataset.

2.6 Solution mapping and the surrogate dataset

In most realistic scenarios, f is a complex computer code that numerically evaluates a governing set of nonlinear (differential) equations. In some settings, users may only interact

with f in a "black-box" fashion by providing inputs as queries and receiving corresponding outputs as answers. These difficulties make reliable computation of consistency and prediction essentially intractable due to the inability to guarantee globally optimal solutions. Moreover, these issues are exacerbated by the large dimension of the uncertain parameter space, a feature that is quite common in modern science and a manifestation of Richard Bellman's famous coinage – "the curse of dimensionality" [11].

B2BDC handles these challenges by combining the precepts of solution mapping [44, 46, 18, 17] with techniques from polynomial optimization [90, 67]. The core idea of solution mapping is to use simple algebraic response surfaces $\{M_i\}_{i=1}^N$ – usually low-order polynomials – as surrogates for the complicated mappings $\{f_i\}_{i=1}^N$ in the computation. To alleviate the crisis posed by dimensionality, the surrogates are constructed independently on a per QOI basis and over only a subset of the n parameters – the *active variables*, i.e., those parameters deemed most impactful to the QOI response. Of course, the choice of active variables may be different for different QOIs. To emphasize the necessity of such a strategy, consider the following. In [42], Frenklach and collaborators investigate a combustion system consisting of $n = 102$ uncertain parameters. A simple sampling strategy accounting for two coordinate perturbations in each parameter dimension would require 2^{102} evaluations of f. Clearly, a different strategy is needed to explore the behavior of QOIs over the parameter space.

Throughout this thesis we will denote the collection of active variable indices for the ith QOI by the ordered subset $\mathcal{A}_i \subset \{1, ..., n\}$. Notationally, we express the active parameters $x_{\mathcal{A}_i} \in \mathbb{R}^{|\mathcal{A}_i|}$ as a subvector of x with entries in $\mathcal{A}_i$, where $|\cdot|$ denotes cardinality. The determination of the active index set $\mathcal{A}_i$ is usually accomplished through a careful sensitivity screening of the inputs to f_i [17, 18, 101, 45]. This screening requires performing an initial *design of computer experiments* over the full prior (in all n parameters), i.e., evaluating $f_i(x)$ on a collection of points in $\mathcal{H}$. Once the active variables are identified, the response surfaces M_i can then be developed by performing a secondary design in the corresponding active variable space, i.e., evaluating $f_i(x)$ at a selection of active inputs (with the inactive components of x fixed at nominal values). Surrogate models can then be fit to the resulting input-output simulation data. Typically the initial screening design exhausts relatively few runs of the computational model. If the active variables make up only a small subset of the full parameters, then developing the surrogate in the active space requires a secondary design with fewer points than what would have been originally required in the full space. The number of computer runs is therefore decreased.

With the computational models replaced by their surrogates, the dataset becomes a collection of polynomial constraints. As a result, Equation (2.4) and Equation (2.5) become polynomial optimization problems. In general, these problems are nonconvex and by no means easy, but their global optima can be reliably bracketed through convex relaxation techniques [47, 67, 90]. We emphasize that the choice of surrogate models in B2BDC is dictated by the ability to make such guarantees. This is further discussed in Section 2.7 where the focus is on quadratic surrogate models, which form the bedrock for most B2BDC applications.

The use of surrogate models is a necessary evil for pragmatic computation and introduces

fitting errors into the analysis:

$$e_i(x) = f_i(x) - M_i(x_{\mathcal{A}_i}). \tag{2.8}$$

The feasible set in Equation (2.3) can therefore be rewritten as,

$$\mathcal{F} = \{x \in \mathbb{R}^n \ : \ x \in \mathcal{H},\ L_i \leq M_i(x_{\mathcal{A}_i}) + e_i(x) \leq U_i \ \text{ for } i = 1, ..., N\}. \tag{2.9}$$

The following approach is often taken to compensate for the error term. First, an estimate $\hat{e}_i^{\max}$ of $\max_{x\in\mathcal{H}} |e_i(x)|$ is computed. In practice, the maximum error is usually only evaluated in the active variables space during the surrogate fitting step. Second, an approximation to $\mathcal{F}$ is developed by widening the bounds to accommodate the error,

$$\mathcal{F}^e = \{x \in \mathbb{R}^n \ : \ x \in \mathcal{H},\ L_i - \hat{e}_i^{\max} \leq M_i(x_{\mathcal{A}_i}) \leq U_i + \hat{e}_i^{\max} \ \text{ for } i = 1, ..., N\}. \tag{2.10}$$

If $\hat{e}_i^{\max} \geq \max_{x\in\mathcal{H}} |e_i(x)|$, then it can be guaranteed that $\mathcal{F}^e$ absorbs $\mathcal{F}$, i.e., $\mathcal{F} \subseteq \mathcal{F}^e$. Ideally, the fitting error for each QOI should be sufficiently small to ensure that the approximation faithfully captures the computational model. As most practical applications of B2BDC necessitate surrogate modeling, the use of $\mathcal{F}^e$ rather than $\mathcal{F}$ is implied. To simplify the notation, we will not distinguish between $\mathcal{F}$ and $\mathcal{F}^e$.

The utility of the solution mapping strategy hinges on the following question: under what conditions can we expect such approximations to be successful? More specifically, when can we expect the response of f to be relatively "benign"? In fields or applications where considerable prior knowledge exists, e.g., $\mathcal{H}$ is relatively small and leads to manageable variation in f, simple approximations such as quadratics may be reasonable. For example, if direct measurement of an uncertain parameter is practical, then the corresponding dimension of $\mathcal{H}$ is likely to be narrow. Similarly, recall that considerable effort goes into developing laboratory experiments, which are often designed to isolate relationships among particular uncertain parameters. In this situation of heterogeneous and semi-isolated QOIs, active variables are likely present and can then be exploited. These points were originally put forth in work by Frenklach et al. [42] to explain the success of the method for a particular combustion chemistry application. Since this initial publication, tools aimed at expanding the applicability of B2BDC techniques have been proposed [36, 45, 97, 34]. These approaches will be discussed in Chapter 3, with an emphasis on a new iterative strategy.

2.7 Computational strategy for NQCQPs

Typical applications of B2BDC use surrogate models of the following forms: polynomial [103], rational quadratic [35], and piecewise quadratic [35, 86]. These modeling choices are specifically motivated by the consistency and prediction formulations of Equation (2.4) and Equation (2.5), which are both optimization problems subject to the dataset constraints. For affine surrogate models, the operations can be effectively addressed by linear programming. Beyond this, however, the computation becomes more complicated. For example, moving

from just affine to quadratic surrogates leads to *nonconvex quadratically constrained quadratic programs* (NQCQPs) [5, 47], a generally NP-Hard class of problems. Determining global solutions becomes intractable as the dimension of the parameter space increases.

Despite these challenges, the above classes of surrogate models lead to optimization problems that admit tractable approximations. Consider a minimization problem $(\mathcal{P})$ with $p^\star$ denoting a global minimum. Then by "tractable approximation", we mean developing two easier-to-solve problems $(\underline{\mathcal{P}})$ and $(\overline{\mathcal{P}})$ with minimums $\underline{p}$ and $\overline{p}$ that bracket the true global minimum, i.e., $\underline{p} \leq p^\star \leq \overline{p}$. In B2BDC, $(\underline{\mathcal{P}})$ is formulated through a convex relaxation of $(\mathcal{P})$ and $(\overline{\mathcal{P}})$ is computed by only asking for a local minimum. The latter task can be accomplished by performing a variant of gradient descent or any nonlinear constrained optimization technique [83, 58]. With the surrogate models described above, $(\underline{\mathcal{P}})$ can be addressed through semidefinite programming. This class of optimization problems is an extension of linear programming in which nonnegativity constraints on vector decision variables are replaced by semidefiniteness constraints on matrix decision variables. Moreover, these problems are convex, meaning local solutions are also global, and efficient algorithms and software are easily accessible [121, 19, 111]. In the general B2BDC setup, formulating and solving semidefinite programs (SDPs) provides the computational engine for generating global guarantees on prediction and inference.

As quadratic surrogate models form the backbone of most B2BDC applications, we review the relaxation techniques for the NQCQP below:

$$\begin{aligned} \min_{z\in\mathbb{R}^n} \quad & q_0(z) \\ \text{s.t.} \quad & L_i \leq q_i(z) \leq U_i \quad \text{for} \quad i = 1,\dots,m \end{aligned} \tag{2.11}$$

where

$$q_k(z) := \begin{bmatrix}1\\ z\end{bmatrix}^\mathsf{T} \mathbf{Q}_k \begin{bmatrix}1\\ z\end{bmatrix} \quad \text{for} \quad k = 0,\dots,m \tag{2.12}$$

and

$$\mathbf{Q}_k = \begin{bmatrix} c_k & b_k^\mathsf{T} \\ b_k & \mathbf{A}_k \end{bmatrix}, \quad \mathbf{A}_k^\mathsf{T} = \mathbf{A}_k. \tag{2.13}$$

The problem in Equation (2.11) is quite general. For example, in the case of computing the SCM in Equation (2.4), the vector $z = \begin{bmatrix} x^\mathsf{T} & \gamma \end{bmatrix}^\mathsf{T}$ and $q_0(z) = -\gamma$ (the optimal solution should then be negated to convert the minimization to a maximization). The number of constraints m could include both the N model-data constraints as well as a set of quadratic constraints defining $\mathcal{H}$. When active variables are present, the matrices $\mathbf{Q}_i \in \mathbb{R}^{(n+1)\times(n+1)}$ may contain different sparsity patterns as dictated by the active indices $\mathcal{A}_i$ identified through sensitivity analysis. There are a multitude of ways to generate or interpret convex relaxations to Equation (2.11). Two methods and an alternate interpretation are described in the subsections below. The following matrix fact is used repeatedly in these derivations.

Theorem 2.7.1. *[102]*

$$\min_x \begin{bmatrix}1\\ x\end{bmatrix}^\mathsf{T} \mathbf{Q} \begin{bmatrix}1\\ x\end{bmatrix} \geq 0 \text{ if and only if } \mathbf{Q} \succeq 0. \tag{2.14}$$

Proof.
$\Leftarrow$: Since $\mathbf{Q} \succeq 0$, $\begin{bmatrix}1\\x\end{bmatrix}^\intercal \mathbf{Q} \begin{bmatrix}1\\x\end{bmatrix} \geq 0$ for all x. Thus $\min_x \begin{bmatrix}1\\x\end{bmatrix}^\intercal \mathbf{Q} \begin{bmatrix}1\\x\end{bmatrix} \geq 0$.
$\Rightarrow$: Take any $v = \begin{bmatrix}v_1\\v_2\end{bmatrix}$ with $v_1 \in \mathbb{R}$. If $v_1 \neq 0$, then

$$\begin{bmatrix}v_1\\v_2\end{bmatrix}^\intercal \mathbf{Q} \begin{bmatrix}v_1\\v_2\end{bmatrix} = v_1^2 \begin{bmatrix}1\\\frac{1}{v_1}v_2\end{bmatrix}^\intercal \mathbf{Q} \begin{bmatrix}1\\\frac{1}{v_1}v_2\end{bmatrix} \geq 0$$

where the last step is justified by the fact that $v_1^2 \geq 0$. If $v_1 = 0$, then

$$\begin{bmatrix}0\\v_2\end{bmatrix}^\intercal \mathbf{Q} \begin{bmatrix}0\\v_2\end{bmatrix} = \lim_{\epsilon \to 0} \begin{bmatrix}\epsilon\\v_2\end{bmatrix}^\intercal \mathbf{Q} \begin{bmatrix}\epsilon\\v_2\end{bmatrix} \geq 0.$$

The first equality is justified by continuity of the quadratic and the second follows since the expression inside the limit is nonnegative for all choices of $\epsilon \neq 0$. □

Lagrangian duality

The problem in Equation (2.11) can be rewritten as

$$p^\star = \min_{z \in \mathbb{R}^n} \max_{\substack{\lambda^l \geq 0 \\ \lambda^u \geq 0}} \mathcal{L}(z, \lambda^l, \lambda^u), \tag{2.15}$$

where λ^l, λ^u are Lagrange multipliers (or dual variables) and $\mathcal{L}(z, \lambda^l, \lambda^u)$ is the Lagrangian,

$$\mathcal{L}(z, \lambda^l, \lambda^u) = q_0(z) + \sum_{i=1}^m \left[\lambda_i^l (L_i - q_i(z)) + \lambda_i^u (q_i(z) - U_i)\right]. \tag{2.16}$$

Interchanging the order of minimization and maximization in Equation (2.15) produces the lower bound,

$$p^\star \geq \underline{p}_{\mathcal{L}} := \max_{\substack{\lambda^l \geq 0 \\ \lambda^u \geq 0}} \min_{z \in \mathbb{R}^n} \mathcal{L}(z, \lambda^l, \lambda^u). \tag{2.17}$$

Note that this formulation is exactly equivalent to

$$\begin{aligned} \underline{p}_{\mathcal{L}} = \max_{\substack{\lambda^l \geq 0 \\ \lambda^u \geq 0 \\ \theta}} \quad & \theta \\ \text{s.t.} \quad & \min_{z \in \mathbb{R}^n} \mathcal{L}(z, \lambda^l, \lambda^u) \geq \theta, \end{aligned} \tag{2.18}$$

where θ is an additional decision variable. Plugging in the expressions for Equation (2.12) into the constraint in a Equation (2.18),

$$\begin{aligned} 0 &\leq -\theta + \min_{z \in \mathbb{R}^n} \mathcal{L}(z, \lambda^l, \lambda^u) \\ &= \min_{z \in \mathbb{R}^n} \begin{bmatrix}1\\z\end{bmatrix}^\intercal \left(\begin{bmatrix} c_0 - \theta & b_0^\intercal \\ b_0 & \mathbf{A}_0 \end{bmatrix} + \sum_{i=1}^m \lambda_i^l \begin{bmatrix} L_i - c_i & -b_i^\intercal \\ -b_i & -\mathbf{A}_i \end{bmatrix} + \sum_{i=1}^m \lambda_i^u \begin{bmatrix} c_i - U_i & b_i^\intercal \\ b_i & \mathbf{A}_i \end{bmatrix} \right) \begin{bmatrix}1\\z\end{bmatrix}. \end{aligned} \tag{2.19}$$

Thus Equation (2.18) can be expressed as the SDP,

$$\begin{aligned} \underline{p}_{\mathcal{L}} = \max_{\substack{\lambda^l \geq 0 \\ \lambda^u \geq 0 \\ \theta}} \quad & \theta \\ \text{s.t.} \quad & \begin{bmatrix} c_0 - \theta & b_0^\intercal \\ b_0 & \mathbf{A}_0 \end{bmatrix} + \sum_{i=1}^{m} \lambda_i^l \begin{bmatrix} L_i - c_i & -b_i^\intercal \\ -b_i & -\mathbf{A}_i \end{bmatrix} + \sum_{i=1}^{m} \lambda_i^u \begin{bmatrix} c_i - U_i & b_i^\intercal \\ b_i & \mathbf{A}_i \end{bmatrix} \succeq 0 \end{aligned} \tag{2.20}$$

where the minimization in the constraint has been replaced with an equivalent linear matrix inequality (LMI), as justified by Theorem 2.7.1. In addition to providing a lower bound on $p^\star$, the Lagrange multipliers λ^l and λ^u provide an additional piece of information. This aspect is revisited in Chapter 6 in the context of the SCM.

Semidefinite relaxation

The semidefinite relaxation approach is motivated by the fact that the quadratics in Equation (2.13) have an equivalent representation,

$$\begin{bmatrix} 1 \\ z \end{bmatrix}^\intercal \mathbf{Q}_k \begin{bmatrix} 1 \\ z \end{bmatrix} = \operatorname{trace}(\mathbf{Q}_k \mathbf{Z}) \quad \text{where} \quad \mathbf{Z} = \begin{bmatrix} 1 \\ z \end{bmatrix} \begin{bmatrix} 1 \\ z \end{bmatrix}^\intercal. \tag{2.21}$$

As such, having z as a decision variable in Equation (2.11) is identical to searching for $\mathbf{Z} \in \mathcal{S}^{1+n}$, where $\mathcal{S}^{1+n} \subseteq \mathbb{R}^{(1+n)\times(1+n)}$ is the subset of symmetric matrices, satisfying the additional criteria

$$\begin{cases} \mathbf{Z}_{11} = 1 \\ \mathbf{Z} \succeq 0 \\ \operatorname{rank}(\mathbf{Z}) = 1. \end{cases} \tag{2.22}$$

Therefore Equation (2.11) can be rewritten with the new matrix variable $\mathbf{Z}$ as

$$\begin{aligned} p^\star = \min_{\mathbf{Z} \in \mathcal{S}^{1+n}} \quad & \operatorname{trace}(\mathbf{Q}_0 \mathbf{Z}) \\ \text{s.t.} \quad & L_i \leq \operatorname{trace}(\mathbf{Q}_i \mathbf{Z}) \leq U_i \quad \text{for} \quad i = 1, \ldots, m \\ & \mathbf{Z}_{11} = 1 \\ & \mathbf{Z} \succeq 0 \\ & \operatorname{rank}(\mathbf{Z}) = 1. \end{aligned} \tag{2.23}$$

In this form, the source of the nonconvexity is isolated in the rank constraint. Eliminating this constraint enlarges the search space for the decision variable $\mathbf{Z}$, thus producing a lower

bound on the minimization as described in Equation (2.24).

$$\begin{aligned} p^\star \geq \underline{p}_{\text{RR}} := \min_{\mathbf{Z}\in\mathcal{S}^{1+n}} \quad & \operatorname{trace}(\mathbf{Q}_0\mathbf{Z}) \\ \text{s.t.} \quad & \operatorname{trace}(\mathbf{Q}_i\mathbf{Z}) \leq U_i \quad \text{for} \quad i=1,\dots,m \\ & -\operatorname{trace}(\mathbf{Q}_i\mathbf{Z}) \leq -L_i \quad \text{for} \quad i=1,\dots,m \\ & \mathbf{Z}_{11} = 1 \\ & \mathbf{Z} \succeq 0. \end{aligned} \tag{2.24}$$

The above relaxation is a semidefinite program in primal form [19]. Moreover, it can be shown that this rank relaxation approach and the Lagrangian approach form a primal-dual pair with $\underline{p}_{\text{RR}} \geq \underline{p}_{\mathcal{L}}$. Under the appropriate constraint qualifications – strict feasibility of the primal problem in Equation (2.24) and a finite solution to the Lagrangian relaxation – strong duality holds and $\underline{p}_{\text{RR}} = \underline{p}_{\mathcal{L}}$ [47, 115]. Note, if the minimizer $\mathbf{Z}^\#$ to Equation (2.24) is rank 1, then

$$\mathbf{Z}^\# = \begin{bmatrix} 1 \\ z^\# \end{bmatrix} \begin{bmatrix} 1 \\ z^\# \end{bmatrix}^\intercal \tag{2.25}$$

for some $z^\# \in \mathbb{R}^n$. By Equation (2.21) and the fact that $\underline{p}_{RR} \leq p^\star$, $z^\#$ is feasible with respect to the original nonconvex problem in Equation (2.11) and therefore a global minimizer.

Stochastic relaxation

The stochastic relaxation approach is a probabilistic reinterpretation of the SDP relaxation. Suppose the decision variable z in Equation (2.11) is a random variable with (finite) first and second moments $\mathbb{E}[z] = \mu_z$ and $\mathbb{E}[zz^\intercal] = \mathbf{\Gamma}_z$. Then,

$$\begin{aligned} \mathbb{E}\left[\begin{bmatrix} 1 \\ z \end{bmatrix}^\intercal \mathbf{Q}_k \begin{bmatrix} 1 \\ z \end{bmatrix}\right] &= \mathbb{E}\left[\operatorname{trace}\left(\mathbf{Q}_k \begin{bmatrix} 1 \\ z \end{bmatrix} \begin{bmatrix} 1 \\ z \end{bmatrix}^\intercal\right)\right] \\ &= \operatorname{trace}\left(\mathbf{Q}_k \begin{bmatrix} 1 & \mu_z^\intercal \\ \mu_z & \mathbf{\Gamma}_z \end{bmatrix}\right) \end{aligned} \tag{2.26}$$

where the last step is due to the linearity of the trace operation. Setting the decision variable $\mathbf{Z}$ in the rank relaxation to

$$\mathbf{Z} = \begin{bmatrix} 1 & \mu_z^\intercal \\ \mu_z & \mathbf{\Gamma}_z \end{bmatrix} \tag{2.27}$$

and noting that $\mathbf{Z} \succeq 0$ is the same as stating that μ_z and $\mathbf{\Gamma}_z$ form a valid moment sequence [13] justifies a stochastic interpretation of Equation (2.24). Hence the rank relaxation in Equation (2.24) may be reexamined as

$$\begin{aligned} \min_{z\in\mathbb{R}^n} \quad & \mathbb{E}\left[q_0(z)\right] \\ \text{s.t.} \quad & L_i \leq \mathbb{E}\left[q_i(z)\right] \leq U_i \quad \text{for} \quad i=1,\dots,m \\ & z \text{ a random vector with first and second moments } (\mu_z, \mathbf{\Gamma}_z). \end{aligned} \tag{2.28}$$

Therefore, the minimizer, denoted earlier by $\mathbf{Z}^{\#}$, produces the first and second moment associated with the random vector z of unknown distribution. This technique suggests a general recipe for computing $\overline{p}$ once $\underline{p}_{RR}$ has been computed [39, 102, 72]: first, construct the covariance matrix $\mathbf{\Sigma}_z = \mathbf{\Gamma}_z - \mu_z \mu_z^\intercal$; second, assume $z \sim N(\mu_z, \mathbf{\Sigma}_z)$; third, sample realizations of z; and finally, use those realizations to initialize a search for a local minimum of Equation (2.11).

2.8 Implications on prediction and inference

The computational strategy in B2BDC enables formal guarantees on both consistency as well as prediction. Consider the prediction problem described in Equation (2.5). For both the extreme values, the computational strategy produces a bracketing given by

$$\underline{p}_L \leq p_L := \min_{x \in \mathcal{F}} M_P(x) \leq \overline{p}_L \tag{2.29}$$

and

$$\underline{p}_U \leq p_U := \max_{x \in \mathcal{F}} M_P(x) \leq \overline{p}_U. \tag{2.30}$$

The extent of the prediction $[p_L, p_U]$ is then bounded internally and externally by the following chain of containments,

$$[\overline{p}_L, \underline{p}_U] \subseteq [p_L, p_U] \subseteq [\underline{p}_L, \overline{p}_U], \tag{2.31}$$

where the first and third intervals are referred to as the *inner* and *outer* bounds. These computations guarantee that the actual prediction uncertainty cannot be narrower (more accurate) than the inner bounds and cannot exceed (be worse than) the outer bounds.

In general, provable statements about a dataset are made on the basis of both the inner and outer bounds. For example, the strategy for the SCM in Equation (2.4) produces inner and outer bounds $\underline{C}_D \leq C_D \leq \overline{C}_D$. If $\overline{C}_D < 0$, then the solution of the SDP proves the dataset is inconsistent. If $\underline{C}_D \geq 0$, the the dataset is verifiably consistent and the local solution also provides a feasible x.

The convention for calling the solution of the convex relaxation an "outer" bound, even for non-prediction quantities such as the SCM, stems from the following fact. The SDP relaxation can be interpreted as resulting from a convex outer approximation to the feasible set. For generality, consider the set defined as follows,

$$\mathcal{G} = \left\{ z \in \mathbb{R}^n : \begin{bmatrix} 1 \\ z \end{bmatrix}^\intercal \mathbf{Q}_k \begin{bmatrix} 1 \\ z \end{bmatrix} \leq 0, \ k = 1, ..., m \right\} \tag{2.32}$$

and note that the SDP relaxation technique described in Section 2.7 produces the following

$$\mathcal{G} \subseteq \widetilde{\mathcal{G}} := \left\{ z \in \mathbb{R}^n : \exists \mathbf{Z} \in \mathcal{S}^{1+n} \text{ such that } \begin{cases} \mathbf{Z}(2{:}1{+}n, 1) = z \\ \mathbf{Z}(1,1) = 1 \\ \mathbf{Z} \succeq 0 \\ \text{trace}(\mathbf{Q}_k \mathbf{Z}) \leq 0 \text{ for } k = 1, ..., m \end{cases} \right\} \tag{2.33}$$

where the submatrix indexing notation $\mathbf{Z}(2{:}1{+}n, 1)$ refers to the 2nd through $(1{+}n)$th entry of the first column. To verify the inclusion, take any $z \in \mathcal{G}$ and construct $\mathbf{Z} = \begin{bmatrix}1\\ z\end{bmatrix}\begin{bmatrix}1\\ z\end{bmatrix}^{\intercal}$.

2.9 A two parameter example

Consider the simple two parameter system from [41] described by the linear ordinary differential equations below,

$$\begin{aligned}\dot{a}(t) &= -k_1 a(t)\\ \dot{b}(t) &= k_1 a(t) - k_2 b(t)\\ \dot{c}(t) &= k_2 b(t)\\ a(0) &= 1,\ b(0) = 0,\ c(0) = 0,\end{aligned} \tag{2.34}$$

where the dot notation $\dot{\square}$ refers to the time derivative. The state variables a,b, and c correspond to concentrations of certain compounds and the uncertain parameter (k_1, k_2) are rates. Let the prior for the uncertain parameters (k_1, k_2) be $\mathcal{H} = [1, 4] \times [0.5, 2]$, and suppose the experimental observations are measurements of the peak response of $b(t)$, the peak time of $b(t)$, and the half time of $a(t)$ as evaluated at some hidden $(k_1^\star, k_2^\star) \in \mathcal{H}$. These $N = 3$ QOIs and the associated uncertainty bounds are shown below in Figure 2.1.

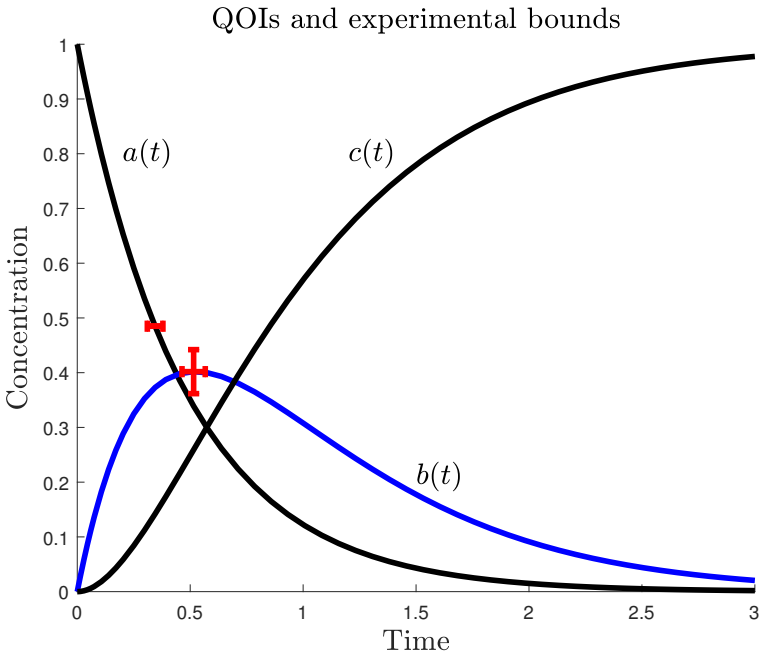

Figure 2.1: QOIs and their associated uncertainty bounds, shown in red.

The first step in the analysis is to generate surrogate models M_1, M_2, and M_3 representing the map between the uncertain parameters and QOIs. This is performed by sampling parameter values from $\mathcal{H}$, evaluating Equation (2.34), and then computing the resulting QOIs. The quadratic surrogate models generated by this procedure are shown below in Figure 2.2.

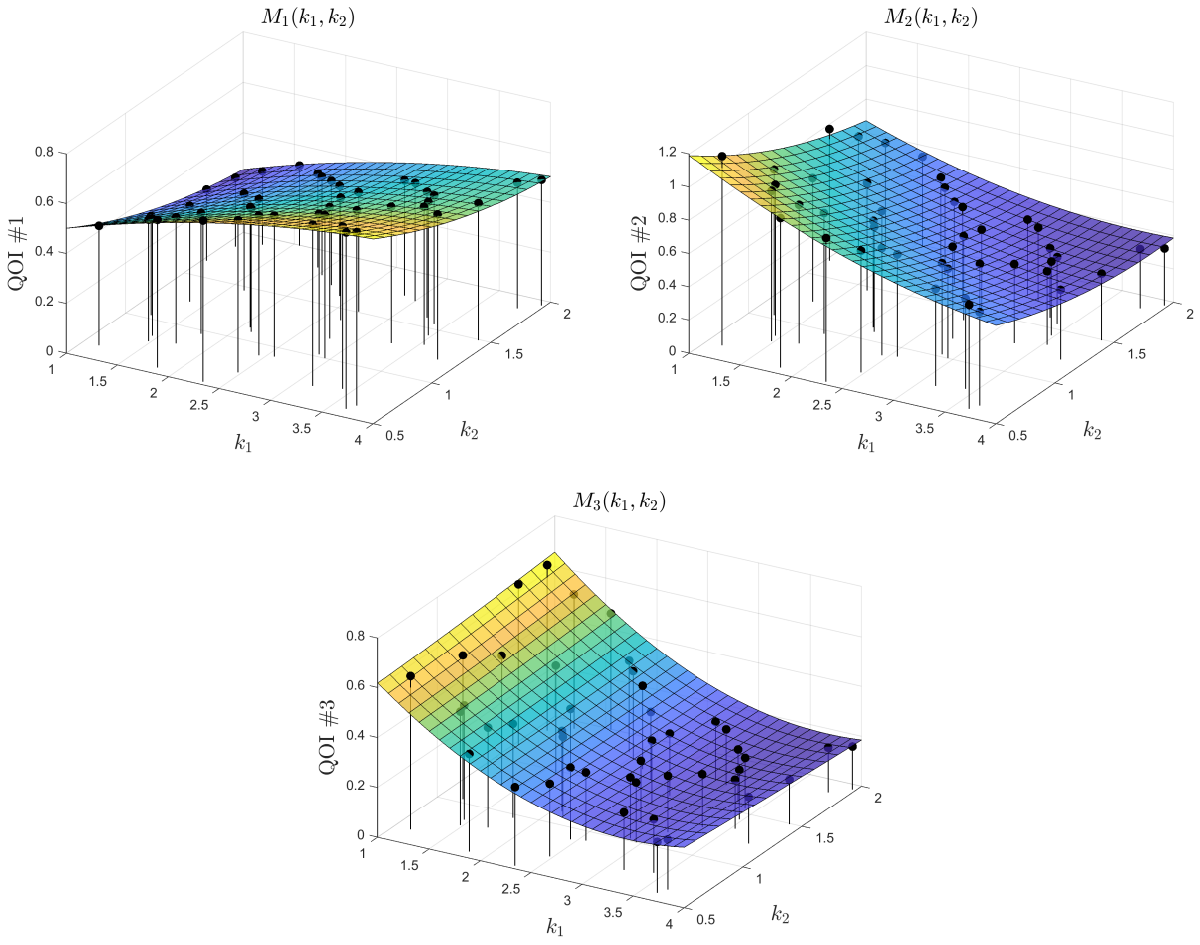

Figure 2.2: Surrogate models with the associated training points shown in black.

The prior, surrogate models, and experimental bounds can then be combined into a dataset. In this two parameter example, the feasible set is easily visualized and, as shown in Figure 2.3, leads to a significant reduction in parameter uncertainty as compared to the prior range. Recall, this reduction is exactly due to the incorporation of observations through the model-data constraints. We note that a key difference between this and more realistic scenarios is the dimensionality. Often, finding feasible samples is not an easy task and computing the scalar consistency measure is essentially required to establish consistency.

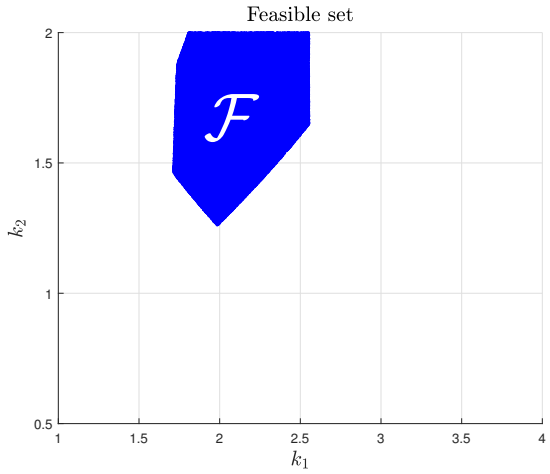

Figure 2.3: The feasible set of the dataset with coordinate axes displaying the full prior range.

Prediction also accounts for this reduction in uncertainty. Let P denote an unmeasured prediction QOI given by the ratio of $c(t_{\max,b})$ to $a(t_{\max,b})$, where $t_{\max,b}$ is the peak time of $b(t)$. The surrogate model corresponding to this new QOI is constructed in the same fashion as the dataset QOIs. As illustrated in Figure 2.4, the inclusion of the experimental data leads to a more certain prediction as compared to using just the prior knowledge.

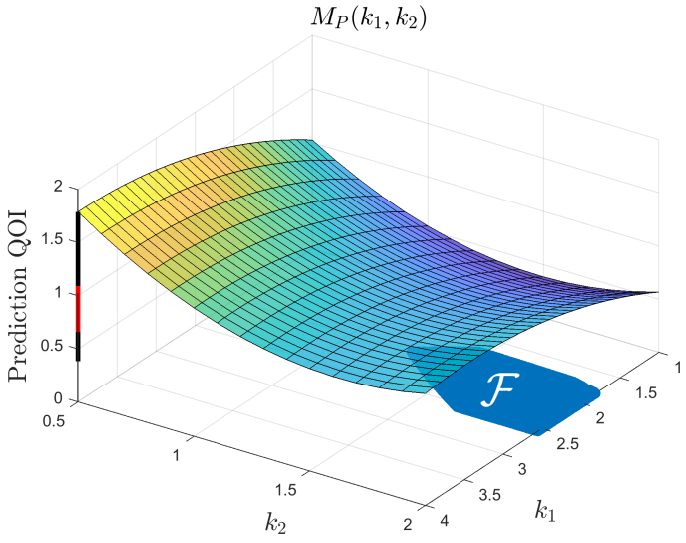

Figure 2.4: Prediction of the surrogate model over the feasible set (in red) and over the prior $\mathcal{H}$ (in black).

2.10 Chapter summary

This chapter reviewed the core B2BDC material, highlighting the fundamental notions of dataset, consistency, and solution mapping. The basic building blocks of the computational strategy for nonconvex quadratically constrained quadratic programs were also discussed. Particular emphasis was placed on the interpretation of the resulting inner and outer bounds, which provide global guarantees on consistency and prediction. While the standard B2BDC material was presented, several of these aspects are expanded upon in subsequent chapters.

Chapter 3

Iterative Construction of a B2BDC Dataset

3.1 Introduction

Reliable uncertainty assessments require accurate surrogate models with well-characterized fitting errors. This aspect was highlighted in the previous chapter. In practice, though, there is always a balance as to what can be achieved. The computational techniques described in Section 2.7 provide a transparent procedure for converting prior and experimental bounds to bounds on prediction. Importantly, the use of convex relaxation and semidefinite programming allows guaranteed statements on both consistency and prediction. The price for these guarantees, however, is a restriction to certain classes of surrogate models: low-order polynomial [42, 103], rational quadratic [35], and piecewise versions of these [35, 86]. In order to use B2BDC in its standard form, all QOIs (of the underlying computational model) must be accurately characterized by the above surrogates over the entirety of $\mathcal{H}$. We include assessment of the fitting error as an aspect of "characterization". This is a limitation of the framework as it may be that certain QOIs are better fit than others.

In this chapter, we draw inspiration from the literature as well as previous extensions of B2BDC to address this challenge. The chapter begins with a brief review of a Bayesian probabilistic approach to uncertainty quantification (UQ) and focuses on the technique of *Bayesian history matching* [28, 29, 116, 117]. We note that B2BDC and Bayesian approaches were compared in previous work by Frenklach and coworkers [41]. The principal conclusions reached by the authors of this study were twofold: first, the "use of both methods protects against possible violations of assumptions in the [Bayesian calibration] approach and conservative specifications and predictions using [B2BDC]"; and second that "[s]hortcomings in the reliability and knowledge of the experimental data can be a more significant factor in interpretation of results than differences between the methods of analysis." The connection between B2BDC and Bayesian history matching is new and was first remarked upon in our recent work [54]. We expand on these links below, demonstrating how the formalism

of history matching relates to previous iteration-based B2BDC extensions and inspires an alternate framework dataset construction.

3.2 Bayesian UQ and history matching

Bayesian calibration

Much of the UQ literature is based on the foundational work of Kennedy and O'Hagan [63], with applications and extensions in, for example, studies by Higdon et al. [56, 55] and Bayarri et al. [9, 10]. A simplified viewpoint of this perspective is presented below.

The general starting point is a probabilistic description of how the model and reality interact. Specifically, the observed data $y_D \in \mathbb{R}^N$ is formulated as the sum of some real world process $y_{\mathcal{R}} \in \mathbb{R}^N$ and a random measurement error $\epsilon \in \mathbb{R}^N$. This random error is generally assumed to have zero mean and to be independent of $y_{\mathcal{R}}$. The real world process is itself formulated as a sum of the mathematical model f evaluated at some unknown best fit parameter $x^\star$ and a bias or model discrepancy $\delta \in \mathbb{R}^N$.

$$\begin{aligned} y_D &= y_{\mathcal{R}} + \epsilon \\ &= f(x^\star) + \delta + \epsilon \end{aligned} \tag{3.1}$$

In a typical Bayesian setting, the uncertain parameter vector x and the discrepancy δ are unknown quantities and represented as random vectors, e.g., with joint density $\pi(x, \delta)$ chosen to reflect available prior information. By assigning a particular distribution for the measurement error ϵ, say Gaussian with independent entries and known variance $\epsilon_i \sim \mathcal{N}(0, \sigma_i^2)$, the likelihood associated with Equation (3.1) becomes,

$$\pi(y_D|\delta, x) = \prod_{i=1}^{N} N(y_{D,i} - f_i(x) - \delta_i, \sigma_i^2). \tag{3.2}$$

Given observed data $y_D \in \mathbb{R}^N$, inference is centered around computing the posterior density,

$$\pi(x, \delta|y_D) = \frac{\pi(y_D|x, \delta)\pi(x, \delta)}{\int \pi(y_D|x, \delta)\pi(x, \delta)dxd\delta}, \tag{3.3}$$

which is often addressed using tools such as Markov Chain Monte Carlo algorithms [49].

Deterministic treatments of uncertainty, as in B2BDC, and probabilistic approaches, such as the Bayesian perspective described above, have distinct interpretations. As argued in detail by Stark [110], replacing constraints with distributions can add unintended information. With this in mind, it would be inappropriate to claim that one approach entirely subsumes the other. Rather than characterizing B2BDC as Bayesian with a specific choice of distribution, we can interpret this deterministic setup by looking at just the supports of distributions. For simplicity, consider the case where there is no discrepancy, i.e., $\delta = 0$. If

we choose the prior and likelihood such that $\pi(x) > 0$ only for $x \in \mathcal{H}$, $\pi(y_{D,i}|x) > 0$ only for $x \in \mathcal{F}_i$ where $i = 1, ..., N$, and assume conditionally independent measurements, then $\pi(y_D|x) = \prod_{i=1}^{N} \pi(y_{D,i}|x)$. The posterior in Equation (3.3) is nonzero only on the feasible set $\mathcal{F}$, and hence supported exactly on $\mathcal{F}$. If the dataset is inconsistent, then the integral in the denominator of Equation (3.3) evaluates to zero. Similar comparisons have been made in the system identification literature, e.g., as presented in a paper by Ninness and Goodwin [81]. We note that the inclusion of model discrepancy δ in the B2BDC framework is investigated in recent work by Li et al. [71]. An interesting connection with consistency and the developments in Chapter 7 can be found in [71, Section 4].

Bayesian history matching

Bayesian history matching [28, 29, 116, 117] is an iterative methodology rooted in the probabilistic setting described above. The objective, however, is not the identification of the full posterior nor the identification of the hypothesized $x^\star$. Rather, the strategy aims only to identify regions of the model's input space that lead to acceptable agreement with observed data. History matching has found successful application in a number of areas, including flow modeling for oil reservoirs [28, 29], galaxy formation [116, 117], rainfall runoff simulation [50], climate modeling [120], epidemiology [2, 3, 1, 76], and systems biology [118]. A prominent feature of the method is the use of stochastic emulators as surrogates for the expensive computational model f, e.g., as described by Sacks et al. [100].

Bayesian history matching works by eliminating regions of the parameter space through a sequence of *waves*, or iterations, whereby an initial region is progressively narrowed. Each wave sees the generation of new parameter samples in the current region and corresponding evaluations of f on those samples. These evaluations are used to train stochastic emulators for a selected collection of QOIs – specifically, those QOIs which can be accurately represented. An expressive and commonly used emulator for $f_i(x)$ is shown below [116, 118]

$$f_i(x) \approx M_i(x) = \sum_j \beta_{ij} g_{ij}(x_{\mathcal{A}_i}) + u_i(x_{\mathcal{A}_i}) + w_i(x) \tag{3.4}$$

where $x_{\mathcal{A}_i}$ once again denotes the collection of active variables relevant to the ith QOI, β_{ij} are unknown coefficients, $g_{ij}(\cdot)$ are selected deterministic functions, $u_i(\cdot)$ is a typically zero mean Gaussian process, and w_i is a white noise process that is uncorrelated with the other components of the emulator. Note that w_i is intended to capture the effect of the inactive variables. A detailed discussion on stochastic emulation and fitting can be found in the tutorial by O'Hagan [85]. Of key importance here is the fact that the emulator has both a mean $\mathbb{E}[M_i(x)]$ and a variance $\mathrm{Var}(M_i(x))$ at each input x.

The samples associated with the particular iteration are then pruned by computing *implausibility measures*, which compare the emulator output to the observed data while accounting for various sources of stochastic uncertainty, including the emulator itself $M_i(\cdot)$, measurement error ϵ_i, and model discrepancy δ_i. The simplest class of implausibility mea-

sures is the univariate family:

$$I_i(x) = \frac{|\mathbb{E}[M_i(x)] - y_D|}{\sqrt{\text{Var}(M_i(x)) + \text{Var}(\delta_i) + \text{Var}(\epsilon_i)}}, \tag{3.5}$$

where i can range over all N QOIs, but is usually restricted to the QOIs selected for the current iteration. Importantly, parameters x whose expected model output deviate substantially from the data may still have small implausibility if there is large uncertainty (as measured through the variance) in the corresponding emulator prediction, discrepancy, and/or observation. Typically, the emulator has large variance in regions far from the training data. Hence, this feature prevents locations of the parameter space where little is known from being deemed implausible. Details on the interpretation and usage of model discrepancy can be found in work by Goldstein and Rougier [51]. For example, Vernon and coauthors [117] suggest choosing the variance of the component δ_i to reflect a subjective assessment as to how capable an expert believes the model is of matching the data for QOI i. Practically speaking, this acts as a tolerance. Larger values of $\text{Var}(\delta_i)$ decrease the implausibility, implying a willingness to tolerate larger deviations from data.

Thresholding the implausibility, e.g., $I_i(x) < c$ for some $c > 0$, imposes a deterministic constraint on the parameter space and defines a region of non-implausible parameter values. Craig and coauthors [29] advocate the choice of $c = 3$, i.e., $I_i(x) < 3$ based on Pukelsheim's three sigma rule [95]. More detailed implausibility measures, such as less aggressive variants or those that account for known correlations, can be found in [116].

After the new non-implausible region is determined, the iteration is updated. In this way, the analysis is continually refined – the non-implausible region cannot grow with each wave, and in the ideal case, becomes a more accurate representation of where the model and data agree. The strategy can be summarized by the following collection of steps. To start, let $\mathcal{X}^k$ denote the current non-implausible region at the beginning of the kth wave. Note, this region $\mathcal{X}^k$ was established in the preceding $(k-1)$st wave. Additionally, let $\mathcal{Q}^k$ contain the indices of QOIs that can be accurately emulated over $\mathcal{X}^k$, i.e., $\mathcal{Q}^k \subseteq \{1, ..., N\}$. For the kth wave [116, 118]:

1. Construct a design of training samples and model evaluations $\{(x^{(p)}, f(x^{(p)}))\}_{p=1}^m$ characterizing the region $\mathcal{X}^k$. For simple cases, this can be accomplished through rejection with a design of (computer) experiments, e.g. Latin hypercube sampling [75] or a Sobol sequence [109]. In many practical scenarios, more efficient strategies are required. For example, slice sampling is discussed in work by Andrianakis et al. [1].

2. Improve / refit each of the emulators that were deemed accurate in the previous wave (i.e., $\mathcal{Q}^{k-1}$) using the training samples from Step 1.

3. Verify if any additional QOIs can be accurately emulated. If so, append the new QOIs to $\mathcal{Q}^{k-1}$, creating $\mathcal{Q}^k$.

4. Update the expressions for implausibility measures $I_i(x)$ with the new emulators for each $i \in \mathcal{Q}^k$.

5. Define the updated non-implausible region as $\mathcal{X}^{k+1} = \{x \in \mathcal{X}^k : I_i(x) < c, \quad i \in \mathcal{Q}^k\}$. This can be accessed through sampling and emulator evaluations.

6. If (a) $\mathcal{Q}^k = \{1, ..., N\}$ and all of the emulators have sufficiently small variance or (b) $\mathcal{X}^{k+1}$ is empty, terminate the algorithm.

7. Iterate $k \leftarrow k + 1$ and go to step 1.

The above strategy ensures that $\mathcal{X}^{k+1} \subseteq \mathcal{X}^k$. As such, the behavior of the computational model f is likely to be smoother over the updated region and QOIs become easier to emulate. Thus the above strategy ensures that the emulators are accurate on relevant or useful regions in the parameter space.

3.3 B2BDC and Bayesian history matching

There are several notable points of comparison between B2BDC and Bayesian history matching. First, both methods evaluate the agreement between a computational model and empirical data by locating a set – the non-implausible region in history matching and the feasible set in B2BDC – defined by hard constraints involving surrogate models, experimental data, and prior knowledge. Both approaches see the existence of this set as a key indicator of "model fits data". Whereas B2BDC directly formulates prediction within the same framework, history matching is viewed as a useful "precalibration" step [117]. A full Bayesian calibration and prediction exercise can be carried out over the reduced parameter space, i.e., the non-implausible region.

As one might expect, differences in both how uncertainty is characterized and the choice of surrogates lead to differences in interpretation. Additionally, a typical application of B2BDC requires the QOIs of the computational model to be well approximated by algebraic surrogates over the prior region $\mathcal{H}$. In this setting, consistency is a global property of the dataset and establishes the existence of a feasible region. A diagnosis of inconsistency proves that the model and data disagree over the entirety of $\mathcal{H}$. In contrast, history matching allows a more flexible class of surrogate models with a more nuanced characterization of uncertainty. A consequence of this, however, is that the non-implausible region at any iteration is assessed locally and on the basis of a finite number of samples. Since assessments of the implausible region are purely sample based, checking if the non-implausible region $\mathcal{X}^k$ is empty at each iteration presents a challenge. This is exacerbated by the fact that in both approaches, the feasible set and the non-implausible set tend to be extraordinarily small relative to the starting region and with potentially complex geometry [116, 1, 99].

3.4 Previous iterative strategies in B2BDC

As discussed earlier, the standard application of B2BDC described in Chapter 2 is a single-shot endeavor and can fail if the surrogate models are poor representations over the prior region $\mathcal{H}$. In previous works, however, iterative strategies aimed at addressing this challenge have been explored. In [36], Feeley et al. proposed a trust-region algorithm based on identifying, translating, and resizing a small rectangular subregion $\mathcal{H}_{\text{tr}} \subseteq \mathcal{H}$, over which all N QOIs can be accurately characterized by quadratic surrogate models. This approach stems from the same strategy in model optimization using solution mapping [46] and later employed in B2BDC [45, 34]. Moreover, Feeley [35] further explored the construction of piecewise quadratic surrogate models by iteratively splitting $\mathcal{H}$ into smaller hyperrectangles and refitting quadratic surrogate models over each subregion. This was extended by Oreluk in [86, Chapter 4], which incorporated consistency checks to enable early termination of the subdivision for provably inconsistent subdomains (subject to maximum fitting error estimates). In addition to piecewise approaches, Edwards and coauthors [34] proposed an iterative domain restriction strategy based on multiple rounds of response surface fitting over a shrinking domain. Similarly, Garcia [48] implemented an algorithm to compute a bounding hyperrectangle during each iteration by refitting surrogate models and computing Equation (2.6).

In the majority of these approaches, the aim was to hone the surrogate modeling in on regions relevant to the feasible set $\mathcal{F}$. This was generally accomplished by fitting all N QOIs at each iteration and computing $\mathcal{F} = \cap_{i=1}^{N} \mathcal{F}_i$ all at once. In such cases, the effectiveness of the implementation essentially relies on accurately characterizing the maximum fitting error for all dataset QOIs. Underestimating this error could lead to truncation, or even removal, of the feasible region and thus incorrect conclusions.

3.5 Incorporating waves in B2BDC

The formalism of Bayesian history matching provides an alternate framework for iteratively constructing a dataset that prioritizes accurate representation of QOIs. In particular, the dataset construction strategies and tools of B2BDC can be embedded within a wave-based scheme in which only the QOIs accurately fit by the typical B2BDC surrogate models are employed at each iteration. Hence, rather than attempting to characterize the full feasible set in one shot, we construct $\mathcal{F}$ through a sequence of waves, initialized with $\mathcal{F}^1 = \mathcal{H}$. As in Section 3.2, the superscript denotes the wave number. Thus during wave k, the dataset is appended with model-data constraints involving surrogates deemed accurate over the current feasible set $\mathcal{F}^k$. At the conclusion of the wave, the updated dataset defines an also-updated feasible region $\mathcal{F}^{k+1}$ with the property that $\mathcal{F}^{k+1} \subseteq \mathcal{F}^k$. Consistency measures can then be recomputed, thus providing a clear signal to either continue or terminate the analysis. If the dataset is inconsistent, tools for investigating the inconsistency, e.g., those discussed later in Chapter 6 and Chapter 7, can be applied. The workflow below describes this approach for

datasets with quadratic surrogate models.

1. Construct a design of training samples and model evaluations $\{(x^{(p)}, f(x^{(p)}))\}_{p=1}^{m}$ for the region $\mathcal{F}^k$.
2. Improve / refit each of the quadratic surrogates $M_i(x)$ that were deemed accurate in the previous wave (i.e., $i \in \mathcal{Q}^{k-1}$) using the training samples from Step 1.
3. Verify if any additional QOIs can be accurately emulated. If so, append the new QOIs to $\mathcal{Q}^{k-1}$, creating $\mathcal{Q}^k$.
4. Append the QOIs in $\mathcal{Q}^k$ to the dataset, thus forming the updated feasible set $\mathcal{F}^{k+1} = \{x \in \mathcal{F}^k : L_i \leq M_i(x) \leq U_i,\ i \in \mathcal{Q}^k\}$.
5. Compute the scalar consistency measure for the dataset.
6. If (a) $\mathcal{Q}^k = \{1, ..., N\}$ or (b) the dataset is provably inconsistent, terminate the algorithm. If no new QOIs can be incorporated, also terminate.
7. Iterate $k \leftarrow k + 1$ and go to step 1.

Note that during the procedure, constraints are only appended to the dataset and never removed. By retaining old constraints, it is guaranteed that $\mathcal{F}^{k+1} \subseteq \mathcal{F}^k$. This restriction is important as during wave k the new surrogate models are developed on, and therefore only valid over, $\mathcal{F}^k$.

3.6 Example: Calibration of an Aerodynamic Body

In this section, we perform a calibration exercise using an example from MATLAB's System Identification Toolbox™ [74, 73] detailing the modeling of an aerodynamic body. For brevity, we suppress the state and output equations; these, along with detailed information about the problem setup, can be found in [74]. The system has 10 inputs, 5 outputs, and contains 16 uncertain parameters. The 10 inputs correspond to time-dependent signals applied at various actuation points along the body and the 5 outputs are the resulting time-dependent responses. A table documenting the uncertain parameters alongside their prior bounds is provided below in Table 3.1.

Table 3.1: The uncertain parameters and the corresponding prior bounds.

Parameter	**Physical meaning**	**Lower Bound**	**Upper bound**
$F \in \mathbb{R}^4$	Aerodynamic force coefficient	$\begin{bmatrix} 18.328 \\ -8.1125 \\ 34.896 \\ 7.2572 \end{bmatrix}$	$\begin{bmatrix} 21.672 \\ -3.8875 \\ 35.104 \\ 18.743 \end{bmatrix}$
$M \in \mathbb{R}^{11}$	Aerodynamic momentum coefficient	$\begin{bmatrix} -1.0619 \\ 14.109 \\ 2.992 \\ -18.335 \\ -2760.8 \\ -54.627 \\ 7.9077 \\ -213.31 \\ -2834 \\ -38.623 \\ -51.205 \end{bmatrix}$	$\begin{bmatrix} -0.93812 \\ 15.891 \\ 3.008 \\ -13.665 \\ -839.18 \\ -45.373 \\ 38.092 \\ -186.69 \\ -1166 \\ 4.6225 \\ -48.795 \end{bmatrix}$
$C \in \mathbb{R}$	Aerodynamic compensation factor	-10.667	0.66713

The prior $\mathcal{H}$ bounds were selected based on constructing a hyperrectangle around an initial guess, which was provided in [74]. In our implementation, the prior was centered and rescaled to the hypercube $\mathcal{H} = [-1, 1]^{16}$.

Experimental data came in the form of 10 specified input signals and the corresponding 5 output signal measurements. Note that the data was generated from a more sophisticated computational code. Hence, it is unknown as to whether the model being analyzed can reliably reproduce these measurements. Additionally, note that these signals were collected at a sampling rate of 50 Hz.

The input signals and initial conditions for the dynamic equations were assumed completely known. In order to simulate uncertain measurements, the output signals were augmented with error bounds of fixed magnitude (for each signal). QOIs were selected to be the responses at 13 time instances from each of the 5 outputs, leading to a total of $N = 65$ QOIs. These signals and the corresponding QOI bounds are shown below in Figure 3.1 and Figure 3.2.

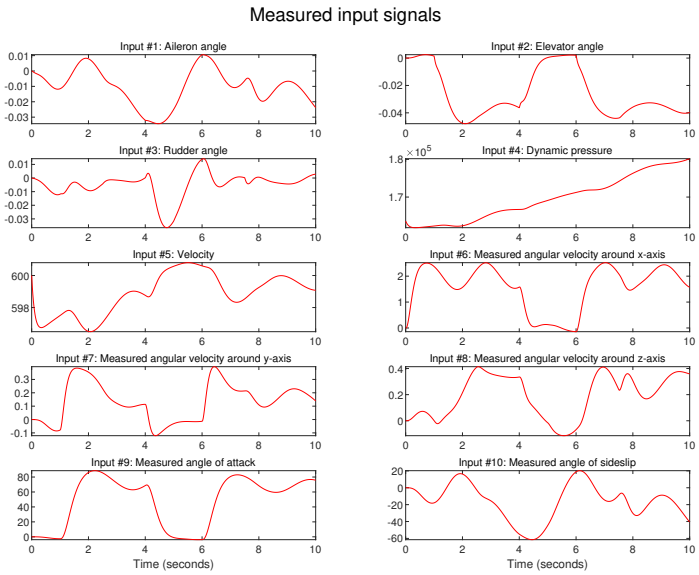

Figure 3.1: Measured input signals.

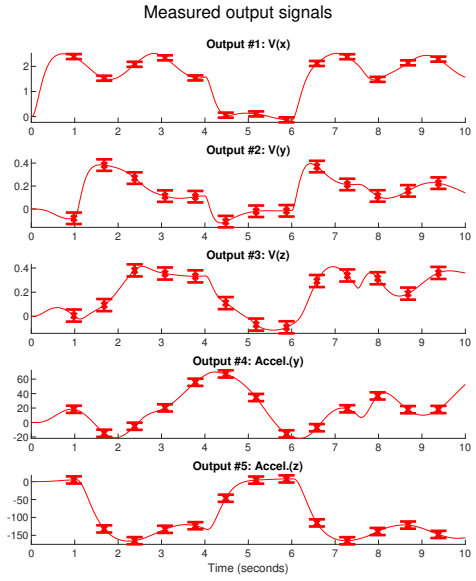

Figure 3.2: Measured output signals. The intervals show the $N = 65$ QOIs and the corresponding uncertainty bounds.

We conducted the iterative B2BDC procedure for a duration of 8 waves, which resulted in the feasible set $\mathcal{F}^9$. During each wave, say wave k, quadratic surrogate models were constructed over $\mathcal{F}^k$. The generation of training samples to fit these surrogates is a non-trivial task. Simple strategies, such as sampling from a bounding box for $\mathcal{F}^k$ using posterior bounds (Equation (2.6)), may often fail. For example, of 5×10^6 Latin hypercube samples generated in the outer bounding box of $\mathcal{F}^9$, none are feasible. This result – i.e., outer bounding and rejection sampling strategies failing to produce feasible samples – is not uncommon [99]. Hence, during each of the $k = 1, ..., 8$ waves 2×10^4 points were generated in the feasible set $\mathcal{F}^k$ using a Gibbs sampling algorithm. Since the sets investigated here are described by quadratic inequalities, the line search step in the sampling setup can be accomplished by root finding. Details on this approach as applied in B2BDC can be found in the forthcoming work by Li et al. [70]. Other sampling strategies for feasible sets in B2BDC or general nonconvex quadratically constrained sets have been investigated in previous work by Russi [97, Chapter 9]. Of the 2×10^4 samples, 7000 were randomly selected to construct the surrogate models.

During the fitting procedure, 50% of the samples were randomly chosen to train the quadratic response surfaces using the supremum norm strategy in [35, Section 7.2.3.1]. The errors were assessed on the entire collection of points. Note that during each wave, a small fraction of the parameter samples were ignored due to the computational model failing to evaluate. In order to gauge the quality of the fit, we computed the following metric:

$$\begin{aligned} \text{EEF}_i &= \frac{(U_i + \hat{e}_i^{\max}) - (L_i - \hat{e}_i^{\max})}{U_i - L_i} \\ &= 1 + \frac{2\hat{e}_i^{\max}}{U_i - L_i} \end{aligned} \tag{3.6}$$

where $\hat{e}_i^{\max}$ is an estimate of the maximum absolute fitting error of QOI i. Equation (3.6) will be referred to as the error expansion factor (EEF), i.e., the factor by which the experimental bounds must be expanded to accommodate the estimated fitting error. Note, the threshold $\text{EEF}_i \leq \alpha$ (for some $\alpha > 1$) is equivalent to a bound on the maximum absolute error $e_i^{\max} \leq (\alpha - 1)\frac{U_i - L_i}{2}$. In this way, the fitting error acceptability is commensurate with the measurement uncertainty – we tolerate larger fitting errors for QOIs with wider experimental uncertainty.

The EEF criteria at the conclusion of both the first and final waves are shown below in Figure 3.3. In this example, we chose to tolerate quadratic models that satisfy an EEF of up to $\alpha = 1.3$.

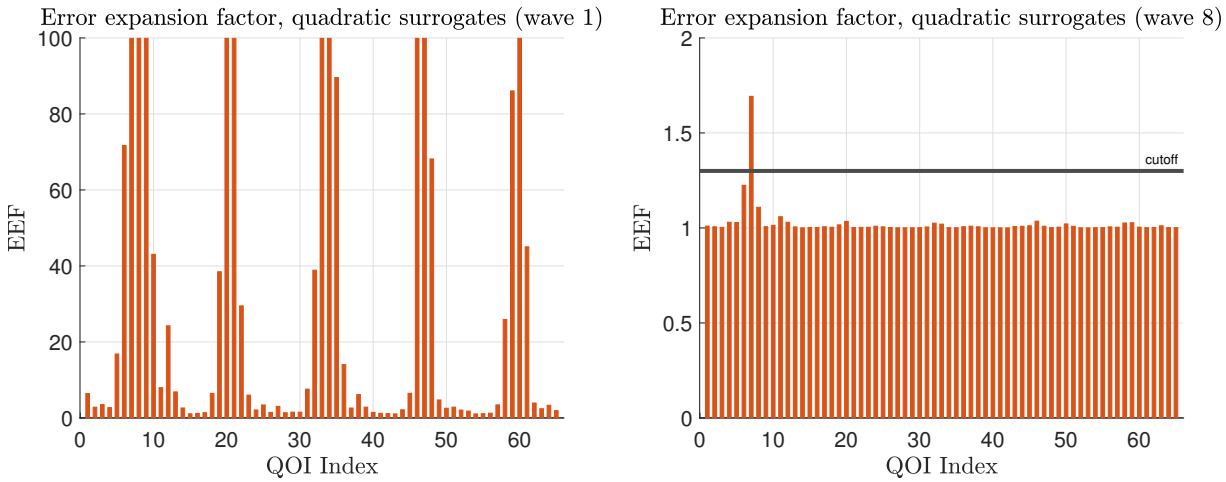

Figure 3.3: Left, EEFs after the first wave. Right, EEFs after the eighth wave.

After eight waves, 64/65 QOIs satisfied the error criteria. The single QOI that did not satisfy the cutoff was not included in the dataset. Hence, QOI #7, which corresponds to the output #1 measurement taken just after $t = 5$ seconds, was not represented through a model-data constraint. The resulting dataset was consistent and the evolution of its content as the waves progressed is tabulated below in Table 3.2. After the sixth wave, the procedure stalls and there is no improvement in the QOIs incorporated. Although the dataset is consistent, the volume of the feasible set becomes negligibly small as compared to the prior.

Table 3.2: Composition of the dataset after the kth wave. The volume fraction is calculated using a collection of 5×10^6 Latin hypercube samples over the prior region $\mathcal{H}$. Note that in each case $\mathcal{F}^{k+1}$ is nonempty.

Wave	$k = 1$	$k = 3$	$k = 5$	$k = 6$	$k = 7$	$k = 8$
# QOIs represented	8	43	56	64	64	64
# total dataset constraints	8	72	179	243	307	371
$\text{vol}(\mathcal{F}^{k+1})/\text{vol}(\mathcal{H})$	0.26	1.8×10^{-4}	2×10^{-7}	≈ 0	≈ 0	≈ 0

The output responses of the underlying computational model for approximately 1000 randomly selected parameter configurations from the prior and the feasible sets at the conclusion of waves $k = 1, 3, 5 - 8$ are shown below in Figure 3.4. An extended vertical axis version of the plot displaying the full variation of the outputs is also shown in Figure 3.5.

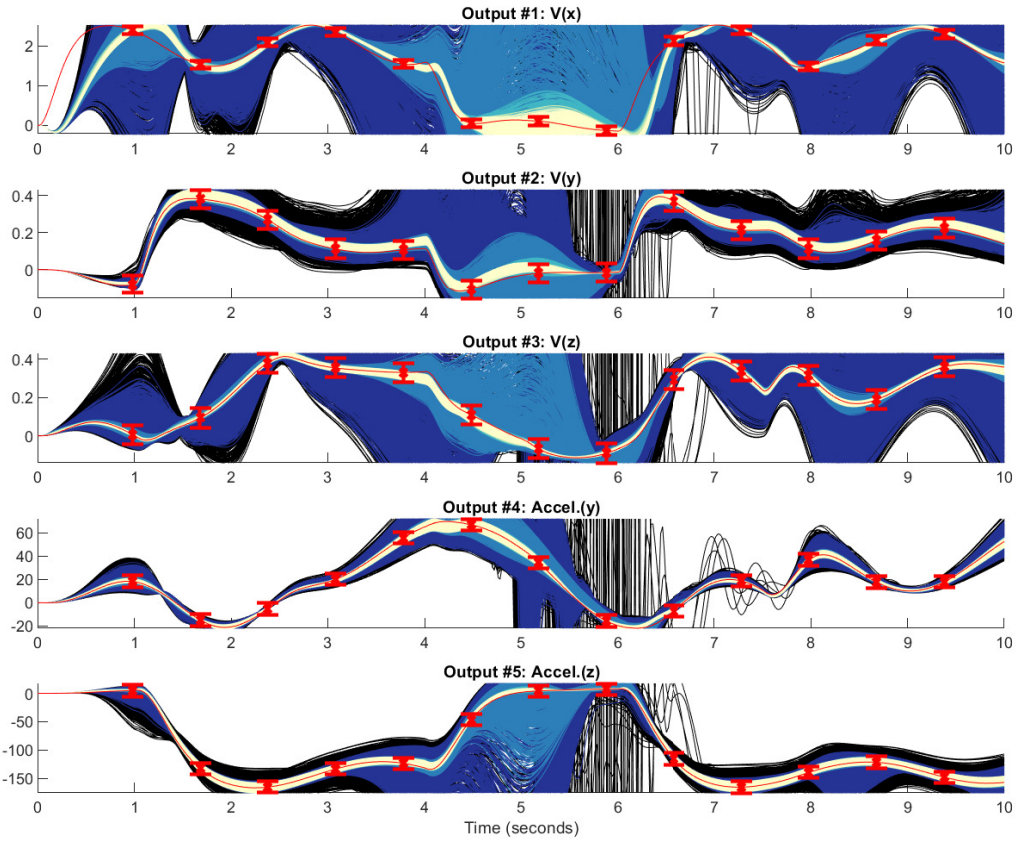

Figure 3.4: Model outputs for approximately 1000 (ignoring failed evaluations) randomly selected feasible parameter vectors after the kth wave. Only the prior and samples after waves $k = 1, 3, 5 - 8$ are shown, with lighter colors corresponding to later waves. Note that wave 8, shown in off-white, is essentially identical to waves 6 and 7. The measured output signals along with the QOI bounds are shown in red.

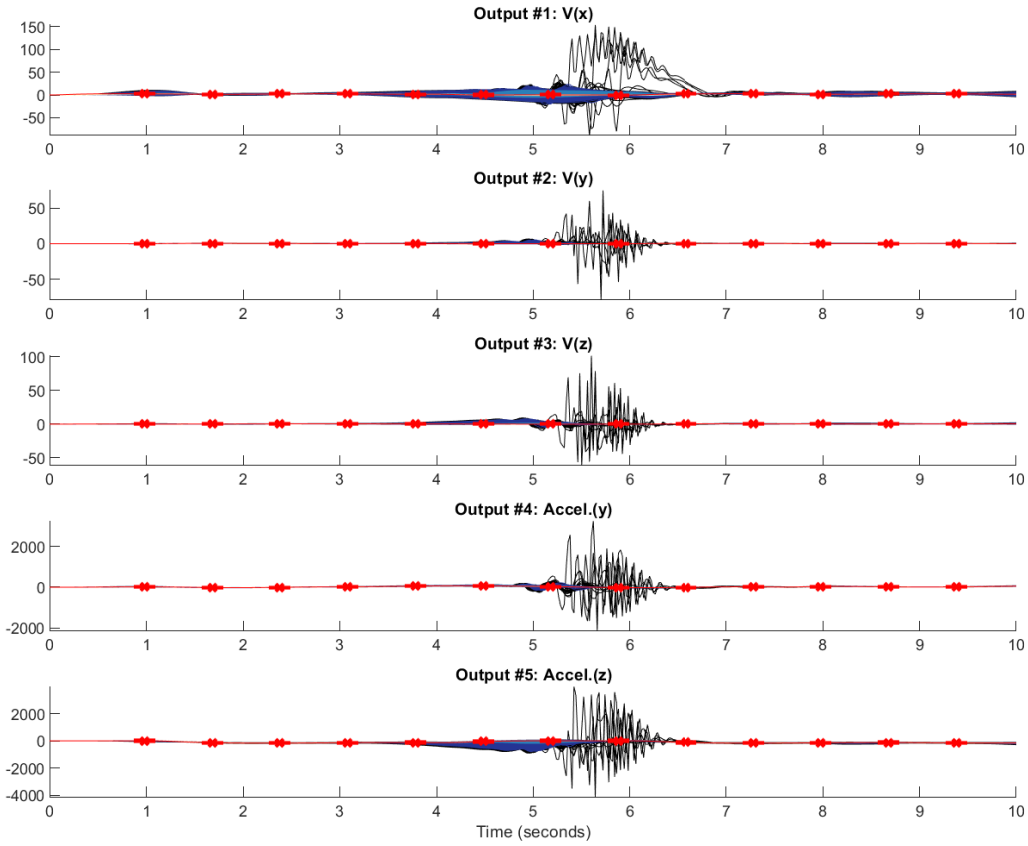

Figure 3.5: Model outputs for 1000 randomly selected feasible parameter vectors after the kth wave. The vertical axis has been extended to accommodate the entire output variation across the feasible parameter samples.

As a general trend, the model outputs after wave $k = 8$, shown as off-white in Figure 3.4, agree well with the experimental bounds. Note, however, the agreement is worst in output 1 and particularly at QOI #7, which was not represented in the dataset due to misfit. We will return to this particular QOI in Chapter 4. For outputs 2 through 5, the feasible samples after wave $k = 8$ adequately capture the behavior of the measured signal at all time instances, i.e., not just the selected QOIs. The first output, however, is quite different. During the time interval $t \in [0, 1)$, significant disagreement exists between the simulated outputs and the measured signal. To investigate this further, we considered a single QOI at $t = 0.18$ seconds fit over the entire prior region, with experimental bounds assessed similar to the other output #1 QOIs. The resulting quadratic surrogate was accurate but inconsistent with the corresponding experimental bounds. Hence, this form of *self-inconsistency* suggests that

the computational model (subject to the choice of $\mathcal{H}$) cannot completely match the measured data over the entire time domain.

In Figure 3.6, we compare a small batch of feasible parameter samples from $\mathcal{F}^9$ (shown in blue) with the best fitting parameter identified using the System Identification Toolbox™ function `nlgreyest` [74], which solves a weighted nonlinear least squares problem involving the computational model and the entire measurement signal. To perform the estimation, we specified the measurement noise variance such that the experimental bounds corresponded to two standard deviations. For lack of a better term, we will refer to the parameter vector identified by the software as "optimal".

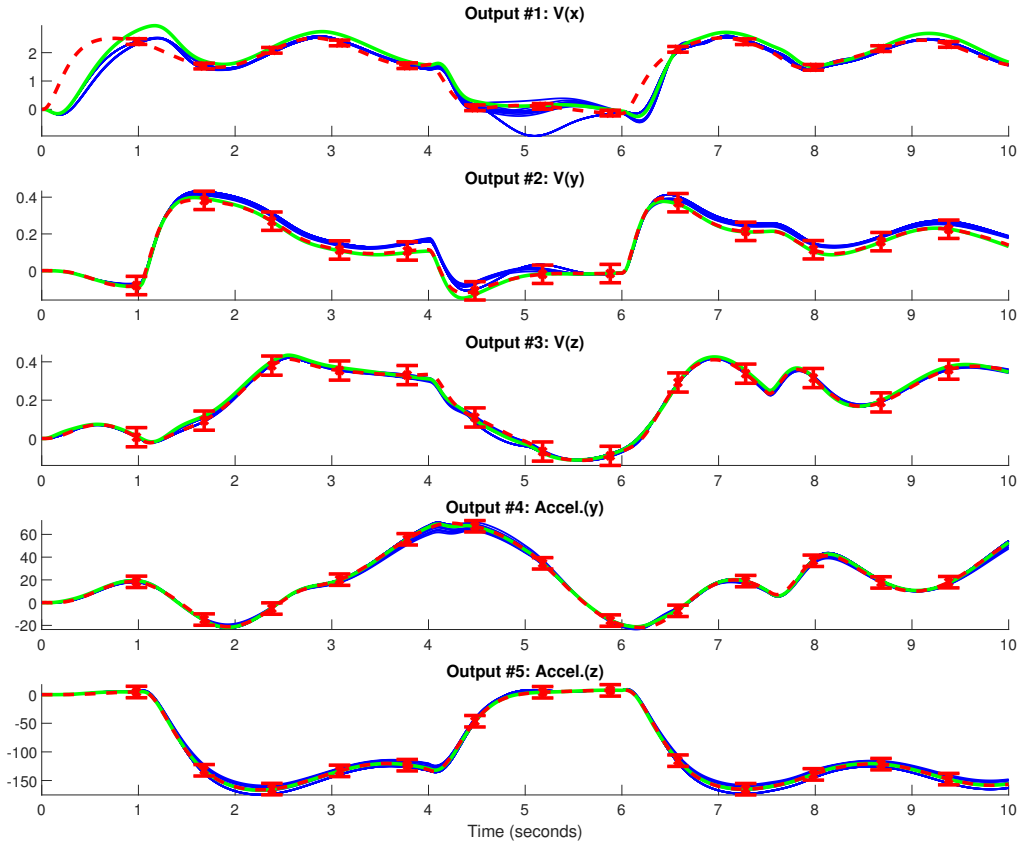

Figure 3.6: The measured output signals are shown in red. 22 feasible samples are shown in blue. The nonlinear least squares optimal point is shown in green.

Notably, the optimal point is not feasible with respect to the dataset. This is evident in the first QOI, where the optimal output signal overshoots the experimental bound. In a sense, the comparison is unfair as the assumptions, settings, and data usage for both

methods are different and the strategies ask different questions. But the core goal remains, to find parameter configurations that lead to acceptable matches between model outputs and measurements. The optimal point tends to outperform the feasible trajectories with respect to the nominal measured signal. When accounting for the uncertainty in the experimental bounds, however, this distinction between optimal and feasible becomes less clear. The majority of the feasible trajectories agree with the experimental bounds and as such should not be discounted from consideration. This is the premise under which B2BDC operates. We note that one of the sampled feasible trajectories significantly undershoots QOI 7 (output #1). This, of course, could be remedied by fitting said QOI with a more sophisticated surrogate model, e.g., a rational quadratic or a piecewise quadratic [35]. The former aspect is further explored in Chapter 4.

Interestingly, the 23 parameter configurations (both feasible and optimal) discussed in the preceding paragraph appear fairly spread out in $\mathcal{H} = [-1, 1]^{16}$. The maximum distance between any two of the displayed feasible parameters is 6.8. For reference, 8 is the maximum distance between any two points in $\mathcal{H}$. The mean and variance of these pairwise distances among the 22 feasible points are 3.13 and 4.83 respectively. The smallest distance between the optimal point and the collection of feasible points is 3.49.

The samples described above provide useful geometric information. Posterior bound calculations, however, confirm the extent of the feasible set. In Figure 3.7, pairwise scatter plots are shown for the feasible samples after the completion of wave $k = 8$ (from Figure 3.4). The corresponding posterior bounds for variables are presented along the horizontal axis.

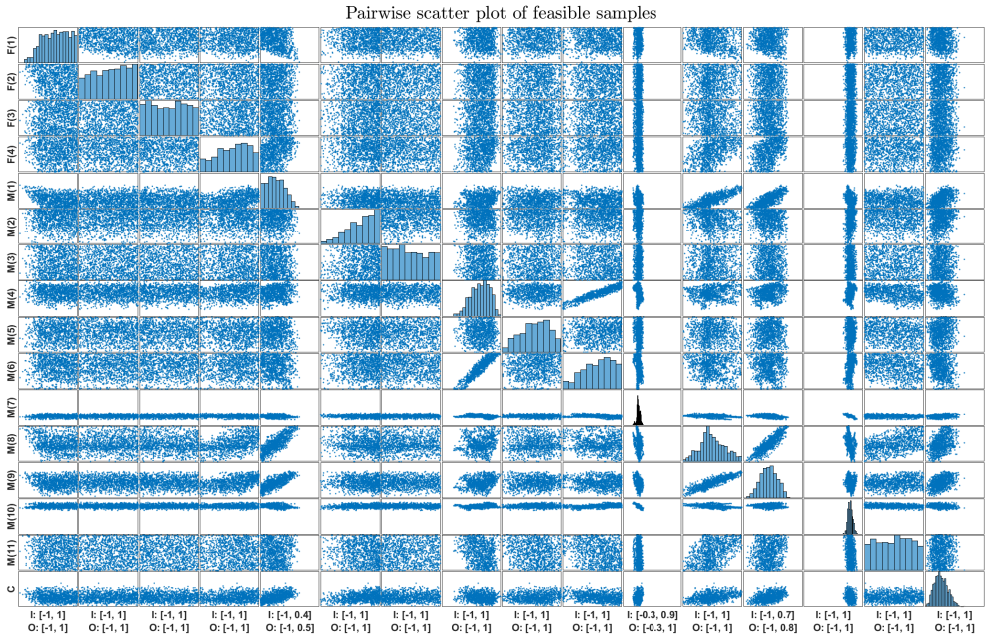

Figure 3.7: Pairwise scatter plot associated with the final batch of feasible samples shown in Figure 3.4. The parameter ordering listed on the vertical axis (top-down) corresponds to the ordering listed on the horizontal axis (left-right). Inner (**I**) and outer (**O**) posterior bounds are shown along the horizontal axis. Note, the axes limits have been standardized to the posterior outer bounds.

Recall, that the inner bounds are computed using local optimization techniques, meaning that the computed bounds are associated with specific extremal feasible parameter vectors. Hence, Figure 3.7 shows that the feasible set verifiably spans the prior range for many of the parameters. For a few, e.g., $M(7)$, the sampled region is only able to cover a fraction of the inner interval, meaning that the samples do not capture the full extent of the feasible region. Interestingly, this general feature of feasible sets that are both long (relative to the prior) and having seemingly minute volume (i.e., compressed in certain dimensions) appears common in many physical examples, e.g., [99].

3.7 Discussion and future work

The above example demonstrate that an iterative, or wave-based, implementation of B2BDC certainly extends the applicability of the technique. If only a few QOIs are accurately fit with quadratic surrogates, there is still the prospect that other QOIs can be

eventually incorporated as the waves progress. In circumstances where this isn't possible, i.e., after several iterations the surrogates cannot be improved, other strategies such as rational quadratic surrogates [35] can be applied. The use of rational quadratics is detailed in the following chapter.

An issue that hasn't been addressed above, however, is the selection of error criteria – both in terms of the metric as well as the cutoff. The choice of the cutoff is rather important, as larger values allow more QOIs and therefore more constraints into the dataset during each wave, thus leading to faster reduction in the prior region. This comes at the cost of being too aggressive and could prematurely eliminate portions of the feasible region due to intersecting many moderately accurate constraints. A viable analogy is that the error criteria acts akin to a "step size". Although we may be able to conclude consistency or inconsistency if the procedure finds a region with good fits, predictions made on that region (in the B2BDC sense, see Section 2.5) must be treated with caution as the full extent of the feasible set may yet to be explored. In such cases, variations of the resize-and-translate strategies reviewed in Section 3.4 can be useful. Another promising direction would be the synthesis of a wave-based strategy with the piecewise approach discussed in work by Oreluk [86].

Chapter 4

Iterative Construction of a B2BDC Dataset: Rational Quadratic Surrogate Models

4.1 Introduction

The tools and methods in B2BDC were initially developed for problems involving just quadratic surrogate models [42]. As emphasized, the reasoning for this was primarily computational – to allow efficient and pragmatic calculation of guaranteed bounds on prediction uncertainty. In the thesis by Feeley [35], it was shown that the same computational strategy involving convex relaxation and semidefinite programming could be applied to surrogate models formulated as ratios of quadratics. This chapter extends the previous iterative dataset construction to incorporating rational quadratic (RQ) surrogate models.

4.2 Fitting over $\mathcal{H}$

RQ surrogate models take the general form of Equation (4.1).

$$M(x) = \frac{\begin{bmatrix}1\\x\end{bmatrix}^\intercal \mathbf{N} \begin{bmatrix}1\\x\end{bmatrix}}{\begin{bmatrix}1\\x\end{bmatrix}^\intercal \mathbf{D} \begin{bmatrix}1\\x\end{bmatrix}} \tag{4.1}$$

where $\mathbf{N} \in \mathcal{S}^{(1+n)}$ and $\mathbf{D} \in \mathcal{S}^{(1+n)}$. To simplify the notation, we use the following shorthand,

$$N(x) = \begin{bmatrix}1\\x\end{bmatrix}^\intercal \mathbf{N} \begin{bmatrix}1\\x\end{bmatrix} \text{ and } D(x) = \begin{bmatrix}1\\x\end{bmatrix}^\intercal \mathbf{D} \begin{bmatrix}1\\x\end{bmatrix}. \tag{4.2}$$

There are a variety of ways to fit a RQ surrogate to training data. In B2BDC, the strategy proposed by Feeley in [35, Section 7.2.3.3] ensures that denominator is positive and well

conditioned over the entirety of $\mathcal{H}$ by enforcing the following condition:

$$1 \leq D(x) \leq \kappa \quad \forall x \in \mathcal{H}, \tag{4.3}$$

where $\kappa > 1$ is a user-specified fitting parameter that bounds the behavior of the denominator $D(x)$. Note that the key challenge in implementing Equation (4.3) is the $\forall x \in \mathcal{H}$ condition. Such conditions are not relevant to constructing quadratic surrogate models, for which standard least squares or supremum norm fitting strategies can be applied. When using only quadratic or polynomial surrogates, the extension to a wave-based implementation of B2BDC is simple and only requires finite samples from $\mathcal{F}^k$ during the kth wave.

In [35, Section 7.2.3.3], a supremum norm fitting strategy ensuring the condition in Equation (4.3) was developed for rectangular $\mathcal{H}$, i.e.,

$$\mathcal{H} = [l_1, u_1] \times \ldots \times [l_n, u_n]. \tag{4.4}$$

First, recall that the linear inequalities implied by $x \in \mathcal{H}$ can be equivalently expressed in the form of quadratics,

$$l_j \leq x_j \leq u_j \iff -x_j^2 + (u_j + l_j)x_j - u_j l_j = \begin{bmatrix} 1 \\ x \end{bmatrix}^\intercal \mathbf{F}_j \begin{bmatrix} 1 \\ x \end{bmatrix} \geq 0 \quad \text{for } j = 1, ..., n. \tag{4.5}$$

In the above, the $\mathbf{F}_j \in \mathcal{S}^{(1+n)}$ are chosen appropriately to make the equivalence. Consider the following pair of inequalities,

$$\begin{aligned} &\exists \lambda \geq 0 \text{ such that } \begin{bmatrix} 1 \\ x \end{bmatrix}^\intercal \left(\mathbf{D} - \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \sum_{j=1}^{n} \lambda_j \mathbf{F}_j \right) \begin{bmatrix} 1 \\ x \end{bmatrix} \geq 0 \quad \forall x \in \mathbb{R}^n \\ &\exists \tau \geq 0 \text{ such that } \begin{bmatrix} 1 \\ x \end{bmatrix}^\intercal \left(\begin{bmatrix} \kappa & 0 \\ 0 & 0 \end{bmatrix} - \mathbf{D} - \sum_{j=1}^{n} \tau_j \mathbf{F}_j \right) \begin{bmatrix} 1 \\ x \end{bmatrix} \geq 0 \quad \forall x \in \mathbb{R}^n. \end{aligned} \tag{4.6}$$

If these inequalities hold, then for any $x \in H$ the condition in Equation (4.3) also holds. Hence, the inequalities Equation (4.6) are sufficient to ensure the denominator is bounded. The advantage of using this sufficient condition, however, is that it can be implemented with the following positive semidefiniteness constraint and thus lends itself to numerical computation,

$$\begin{aligned} &\exists \lambda \geq 0 \text{ such that } \mathbf{D} - \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \sum_{j=1}^{n} \lambda_j \mathbf{F}_j \succeq 0 \\ &\exists \tau \geq 0 \text{ such that } \begin{bmatrix} \kappa & 0 \\ 0 & 0 \end{bmatrix} - \mathbf{D} - \sum_{j=1}^{n} \tau_j \mathbf{F}_j \succeq 0. \end{aligned} \tag{4.7}$$

Given m samples of the underlying function f, namely $\{(x^{(p)}, f(x^{(p)}))\}_{p=1}^{m}$, the supremum norm fitting criteria seeks to pick $\mathbf{N}$ and $\mathbf{D}$ in accordance with the following optimality

criterion,

$$\begin{aligned}
\min_{\substack{t \\ \mathbf{N}\in\mathcal{S}^{(1+n)} \\ \mathbf{D}\in\mathcal{S}^{(1+n)}}} \quad & t \\
\text{s.t.} \quad & \left|\frac{N(x^{(p)})}{D(x^{(p)})} - f(x^{(p)})\right| \leq t \qquad \text{for } p = 1, ..., m \\
& \mathbf{D} - \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \sum_{j=1}^{n} \lambda_j \mathbf{F}_j \succeq 0 \\
& \begin{bmatrix} \kappa & 0 \\ 0 & 0 \end{bmatrix} - \mathbf{D} - \sum_{j=1}^{n} \tau_j \mathbf{F}_j \succeq 0 \\
& \lambda \geq 0, \ \tau \geq 0,
\end{aligned} \tag{4.8}$$

where the fitting parameter κ is user-specified. Since the denominator of the RQ model is ensured to be positive, we may rewrite Equation (4.8) as,

$$\begin{aligned}
\min_{\substack{t \\ \mathbf{N}\in\mathcal{S}^{(1+n)} \\ \mathbf{D}\in\mathcal{S}^{(1+n)}}} \quad & t \\
\text{s.t.} \quad & -tD(x^{(p)}) \leq N(x^{(p)}) - f(x^{(p)})D(x^{(p)}) \leq tD(x^{(p)}) \qquad \text{for } p = 1, ..., m \\
& \mathbf{D} - \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \sum_{j=1}^{n} \lambda_j \mathbf{F}_j \succeq 0 \\
& \begin{bmatrix} \kappa & 0 \\ 0 & 0 \end{bmatrix} - \mathbf{D} - \sum_{j=1}^{n} \tau_j \mathbf{F}_j \succeq 0 \\
& \lambda \geq 0, \ \tau \geq 0.
\end{aligned} \tag{4.9}$$

If t were fixed, the above problem would be an SDP and therefore could be addressed using typical SDP solvers. The interaction between the optimization variables t and $\mathbf{D}$, however, is an obstacle. The standard approach to handling this, as discussed in [35], is to perform bisection on the scalar variable t. Hence at each iteration of the bisection, we fix a value for t and verify the feasibility of Equation (4.9) by solving an SDP.

4.3 Fitting over $\mathcal{F}$

Consider a dataset composed of RQ surrogate models. Naturally, this includes quadratics as well. There are two immediate ways of extending the previous strategy of fitting over $\mathcal{H}$ to accommodate fitting over $\mathcal{F}$. The first will be referred to as **RQ-H** and directly implements the previous technique to training data from $\mathcal{F}$, i.e., $\{(x^{(p)}, f(x^{(p)}))\}_{p=1}^{m}$ where $x^{(p)} \in \mathcal{F}$ for all $p = 1, ..., m$. This approach can be overly restrictive as it requires controlling the behavior of the denominator over all of $\mathcal{H}$ rather than just $\mathcal{F}$. The basic reasoning here is that for

$\mathcal{F} \subseteq \mathcal{H}$, the requirement $\forall x \in \mathcal{H}$ is more restrictive than $\forall x \in \mathcal{F}$ and therefore needlessly limits the search for admissible coefficient matrices $\mathbf{N}$ and $\mathbf{D}$. The second strategy is to only control the denominator on $\mathcal{F}$. This approach will be referred to as **RQ-F** and is described below. Note that both **RQ-H** and **RQ-F** can be easily embedded within the iteration-based scheme of Chapter 3 by fitting over $\mathcal{F}^k$ during the kth iteration.

For a dataset comprised of RQ surrogate models, the feasible set $\mathcal{F}$ is defined as,

$$\mathcal{F} = \left\{ x \in \mathcal{H} : L_i \leq \frac{N_i(x)}{D_i(x)} \leq U_i, \text{ for } i = 1, ..., N \right\}. \tag{4.10}$$

In this scenario, the rational quadratic models in Equation (4.10) are known and the objective is to construct $N(x)$ and $D(x)$ over $\mathcal{F}$. Additionally, note that this setup also includes quadratic surrogates – just set $D_i(x) = 1$ for the appropriate QOI i. To ensure that the denominator $D(x)$ is well behaved on $\mathcal{F}$, we replace the condition in Equation (4.3) with the following,

$$1 \leq D(x) \leq \kappa \quad \forall x \in \mathcal{F}. \tag{4.11}$$

By definition, $\mathcal{F}$ also includes the constraints defining $\mathcal{H}$. Note that $x \in \mathcal{F}$ if and only if the following inequalities hold for all $i = 1, ..., N$,

$$\begin{aligned} \begin{bmatrix} 1 \\ x \end{bmatrix}^\intercal (\mathbf{N}_i - L_i \mathbf{D}_i) \begin{bmatrix} 1 \\ x \end{bmatrix} &\geq 0 \\ \begin{bmatrix} 1 \\ x \end{bmatrix}^\intercal (U_i \mathbf{D}_i - \mathbf{N}_i) \begin{bmatrix} 1 \\ x \end{bmatrix} &\geq 0. \end{aligned} \tag{4.12}$$

In order to enforce Equation (4.11), the following linear matrix inequality can be imposed, paralleling the development of Equation (4.7),

$$\begin{aligned} &\exists \lambda \geq 0,\ \sigma^L, \sigma^U \geq 0 \text{ such that} \\ &\quad \mathbf{D} - \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \sum_{j=1}^{n} \lambda_j \mathbf{F}_j - \sum_{i=1}^{N} \left(\sigma_i^L (\mathbf{N}_i - L_i \mathbf{D}_i) + \sigma_i^U (U_i \mathbf{D}_i - \mathbf{N}_i) \right) \succeq 0 \\ &\exists \tau \geq 0,\ \nu^L, \nu^U \geq 0 \text{ such that} \\ &\quad \begin{bmatrix} \kappa & 0 \\ 0 & 0 \end{bmatrix} - \mathbf{D} - \sum_{j=1}^{n} \tau_j \mathbf{F}_j - \sum_{i=1}^{N} \left(\nu_i^L (\mathbf{N}_i - L_i \mathbf{D}_i) + \nu_i^U (U_i \mathbf{D}_i - \mathbf{N}_i) \right) \succeq 0. \end{aligned} \tag{4.13}$$

Hence, the fitting can be accomplished for training data $\{(x^{(p)}, f(x^{(p)}))\}_{p=1}^m$ by the following

optimization,

$$
\begin{aligned}
\min_{\substack{t \\ \mathbf{N}\in\mathcal{S}^{(1+n)} \\ \mathbf{D}\in\mathcal{S}^{(1+n)}}} \quad & t \\
\text{s.t.} \quad & -tD(x^{(p)}) \le N(x^{(p)}) - f(x^{(p)})D(x^{(p)}) \le tD(x^{(p)}) \quad \text{for } p = 1, ..., m \\
& \mathbf{D} - \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \sum_{j=1}^{n} \lambda_j \mathbf{F}_j - \sum_{i=1}^{N} \left(\sigma_i^L (\mathbf{N}_i - L_i \mathbf{D}_i) + \sigma_i^U (U_i \mathbf{D}_i - \mathbf{N}_i) \right) \succeq 0 \\
& \begin{bmatrix} \kappa & 0 \\ 0 & 0 \end{bmatrix} - \mathbf{D} - \sum_{j=1}^{n} \tau_j \mathbf{F}_j - \sum_{i=1}^{N} \left(\nu_i^L (\mathbf{N}_i - L_i \mathbf{D}_i) + \nu_i^U (U_i \mathbf{D}_i - \mathbf{N}_i) \right) \succeq 0 \\
& \lambda \ge 0, \ \sigma^L, \sigma^U \ge 0 \\
& \tau \ge 0, \ \nu^L, \nu^U \ge 0.
\end{aligned}
\tag{4.14}
$$

As before, this can be solved by performing bisection on t and solving an SDP feasibility problem at each step.

The above strategy **RQ-F** was motivated by the fact that **RQ-H** could be needlessly restrictive for the objective of fitting over $\mathcal{F}$. To make this precise, suppose fixed $\kappa > 1$ and training data $\{(x^{(p)}, f(x^{(p)}))\}_{k=1}^{m}$ with each $x^{(p)} \in \mathcal{F}$. Let $(t^\star, \mathbf{N}^\star, \mathbf{D}^\star, \lambda^\star, \tau^\star)$ be optimal for **RQ-H**. Set $\sigma^L = \sigma^U = \nu^L = \nu^U = 0$. Then, the collection $(t^\star, \mathbf{N}^\star, \mathbf{D}^\star, \lambda^\star, \sigma^L, \sigma^U, \tau^\star, \nu^L, \nu^U)$ is feasible for **RQ-F**. Therefore, the optimal values of the fitting procedures follow **RQ-F** $\le$ **RQ-H**, implying that fitting over $\mathcal{F}$ cannot produce a worse RQ surrogate than fitting over $\mathcal{H}$. Note that these relationships only apply to the maximum absolute error on the training data. In most applications, we are interested in the error calculation on the test data.

4.4 Sampling RQ datasets

In forthcoming work by Li et al. [70] and work by Russi [97, Chapter 9], general strategies were developed for sampling nonconvex quadratically constrained sets. These tools were used during each iteration in Section 3.6 to generate feasible parameter samples for fitting surrogate models. Due to Equation (4.10) and Equation (4.12), sampling a set defined by N RQ model-data constraints can be reformulated as sampling a set of $2N$ one-sided quadratic inequalities. Hence, an entire waves procedure can be completed for RQ surrogate models with the existing strategies. In the following example, however, we focus on completing the previous example from Section 3.6.

4.5 Example continued: calibration of an aerodynamic body

The calibration of an aerodynamic body example in Section 3.6 demonstrated that an iterative application of B2BDC can be useful. Due to the difficulty in generating an accurate quadratic fit for QOI #7, however, the exercise was essentially left incomplete. The surrogate for this QOI did not satisfy the accuracy criteria and therefore the associated model-data constraint was not accounted for in the dataset. The outcome of this can be seen in Figure 3.4, in which feasible samples failed to match the corresponding experimental bound. Here, we revisit the final wave and compare quadratic, **RQ-H**, and **RQ-F** fits for the missing QOI.

For each of the surrogate types, the same set of training and test data was used. In contrast to the result in Section 3.6, 700 points were set aside for training and the remaining 6300 were used to calculate the errors and EEF criteria. For both **RQ-H** and **RQ-F**, the following values of κ were investigated: $\kappa \in \{2, 2.5, 3, ..., 29.5, 30\}$. For each of these values, an RQ surrogate was fit. The surrogate with the smallest maximum absolute test error was selected as the optimal representation. Histograms of the fitting errors as evaluated on both training and test points are displayed below in Figure 4.1. The error distributions for both **RQ-H** and **RQ-F** are nearly identical and both are significant improvements over the quadratic surrogate.

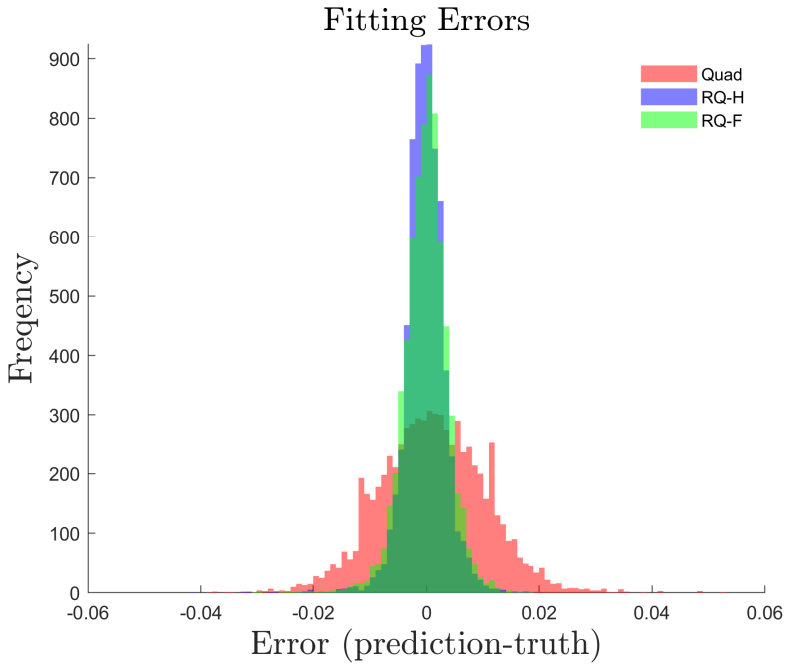

Figure 4.1: Fitting error histograms for each of the surrogate types.

The EEF criteria, which is assessed based on the maximum absolute test error, for all

three surrogate types is plotted below in Figure 4.2.

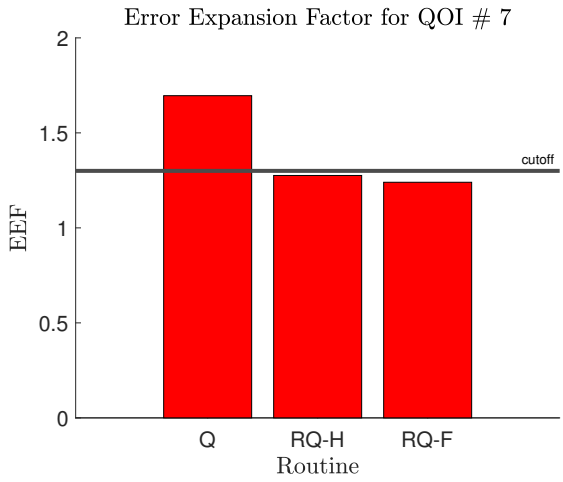

Figure 4.2: EEF criteria for the three types of surrogate models: quadratic, **RQ-H**, and **RQ-F**.

The **RQ-F** surrogate has the lowest EEF as compared to the other three criteria. Both RQ strategies satisfy the cutoff of 1.3. As a result, the model-data constraint with the **RQ-F** surrogate was appended to the dataset. The inclusion of the constraint improves the quality of the feasible set. In Figure 4.3, we re-plot the result in Figure 3.4 with the impact of the new dataset superimposed over the previous model evaluations. Recall that the traces shown in the figure result from the computational model being evaluated on samples judged feasible using the surrogate models.

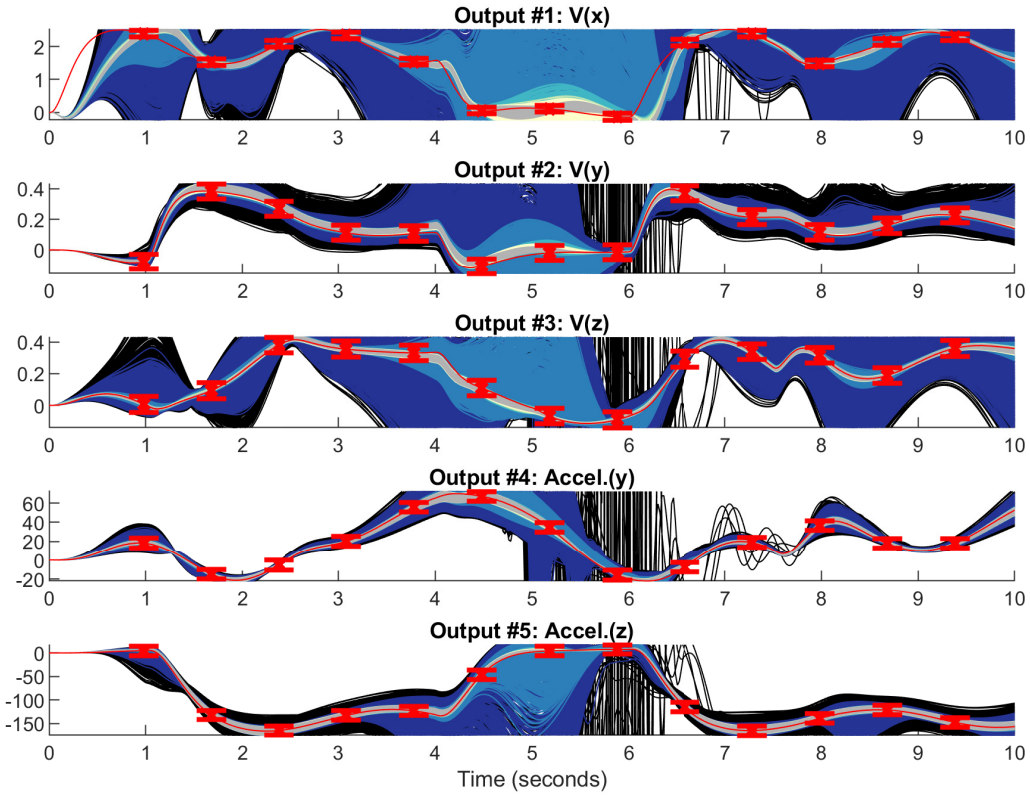

Figure 4.3: Model outputs after the kth wave. The model evaluations shown in gray are feasible with respect to the new dataset, which includes QOI #7 fit with an RQ surrogate.

The agreement between the model traces and the experimental bounds are improved. As in Section 3.6, we evaluate a small batch of feasible parameter samples alongside the optimal parameter vector found using the System Identification Toolbox™ function `nlgreyest` [74]. The resulting plot is shown below in Figure 4.4.

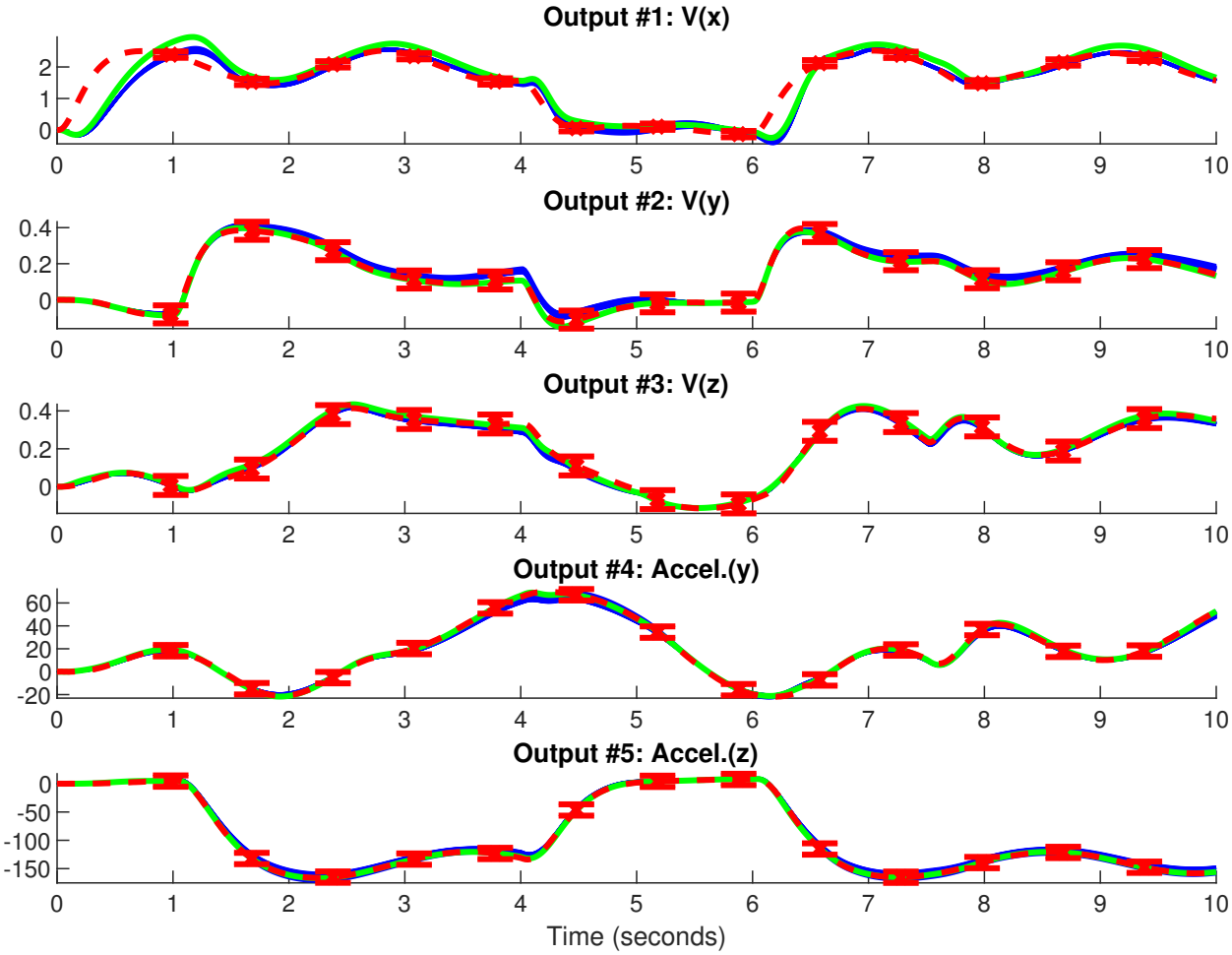

Figure 4.4: The measured output signals are shown in red. Evaluations of 22 feasible samples are shown in blue. The nonlinear least squares optimal point evaluation is shown in green.

Importantly, the inclusion of QOI #7 into the dataset leads to closer agreement with the experimental bounds in output #1. As before, the distinction between the optimal and feasible traces is generally unclear at the designated QOIs. Although the optimal parameter vector better agrees with the nominal measurement, a proper accounting of uncertainty would require the consideration of all feasible parameter vectors.

4.6 Chapter summary

This chapter extended the use of rational quadratic surrogate models to the iterative, wave-based construction of a dataset. In this, we discussed two fitting strategies for developing rational quadratics over a feasible set. The first, **RQ-H**, is essentially identical to the

method proposed by Feeley in [35, Section 7.2.3.3]. The second, **RQ-F**, is a direct extension which exploits the dataset constraints to a greater extent. These strategies were applied to successfully complete the calibration exercise initiated in Chapter 3.

Chapter 5

Additional Feasibility Criteria

5.1 Introduction

The conceptual similarities between B2BDC and Bayesian history matching discussed in Chapter 3 prompt an interesting question: can other aspects from history matching, such as the use of implausibility measures, be adopted into B2BDC in a useful way? Clearly, implausibility is one notion for quantifying model-data agreement. Consistency in B2BDC is another. Although the contexts and implementations for these measures depend on how uncertainty is framed, they both address a similar concept – the existence, location, and isolation of useful regions in the parameter space. These facts motivate an extension of the previous question: can alternate notions of feasibility, in general, be incorporated into a B2BDC dataset?

In this chapter, we discuss techniques that allow sample-based feasibility criteria, such as the implausibility measure, to be incorporated into a B2BDC analysis in an approximate manner. What is important is that the criteria identify and label a finite collection of samples as "feasible". For example, these samples could satisfy a model-data constraint of the general form in Equation (2.2) with a more sophisticated surrogate model outside of those typically used in B2BDC. If this constraint was then included in the dataset, then the standard computational strategy for computing outer bounds and global guarantees would not apply. The primary strategy we investigate here is the use of minimum volume bounding ellipsoids to characterize collections of feasible samples. A similar objective can be reached by employing quadratic kernel support vector machines as well. In essence, these approaches develop geometric approximations of the feasible samples that are encoded through quadratic constraints, and hence cohere with B2BDC's standard computational protocol. These constraints can then be incorporated into a dataset as additional assertions and thus accounted for in the analysis.

5.2 Bounding ellipsoids to feasible samples

The first strategy we discuss is to approximate collections of feasible parameter samples, say with respect to a QOI i, with a bounding ellipsoid of minimal volume. These ellipsoids can then be included in the dataset as additional quadratic constraints, thus indirectly enforcing the feasibility criteria imposed by the samples. We note that ellipsoids have been used in the past as conservative representations of "feasible" regions in system identification [38, 94, 82] and more generally as approximations to data points in n-dimensional space [21, 96].

Our focus on ellipsoidal bounding regions is primarily due to simplicity and computational tractability. Since an ellipsoid can be represented by a single quadratic inequality, the optimization techniques typically used in B2BDC to analyze dataset consistency and prediction can be applied without modification. In general, additional linear constraints or quadratic inequalities can also be included without affecting the existing computational infrastructure. We note that techniques to compute bounding ellipsoids for sets described by general polynomial inequalities have been investigated by others [80, 64]. Similar approaches have also been investigated in B2BDC [99] as outer approximations to existing feasible sets for the purpose of, for example, sampling and volume computation. The approach pursued here, however, differs in both motivation and methodology. The aim is to develop coarse representations of potentially complicated feasible regions to include in the analysis. Although these coarse representations are approximate, further refinements can be made through the iterative approach discussed in Chapter 3. This will be demonstrated by example in Section 5.5. It should be also noted that more sophisticated descriptions of the bounding region can be formulated via polynomial inequality constraints [30, 66, 31, 69]. The scalability of both these techniques and subsequent analyses, however, remains an open question. In contrast, fast and efficient algorithms exist for computing enclosing ellipsoids [112, 114].

5.3 Minimum volume enclosing ellipsoids

Let $\{x^{(p)}\}_{p=1}^m \subseteq \mathbb{R}^n$ be a finite collection of points. These points can be considered samples from the feasible set $\mathcal{F}_i$ of a particular QOI i. In contrast to previous chapters, we make no judgement on how $\mathcal{F}_i$ is defined. Recall, an ellipsoid in $\mathbb{R}^n$ can be described by the quadratic inequality

$$\mathcal{E}(\mathbf{Q}, \bar{x}) = \{x \in \mathbb{R}^n : (x - \bar{x})^\intercal \mathbf{Q}(x - \bar{x}) \leq 1\}, \tag{5.1}$$

where $\bar{x} \in \mathbb{R}^n$ is the center and $\mathbf{Q} \in \mathbb{R}^{n \times n}$, with $\mathbf{Q} \succ 0$, dictates the shape. In order for the "minimum volume" criteria to be sensible, we require that the vectors $\{x^{(p)}\}_{p=1}^m$ not be contained in any hyperplane, i.e., the points must have an affine hull of full dimension n. It is well known that the volume of an ellipse $\mathcal{E}(\mathbf{Q}, \bar{x})$ is inversely proportional to the determinant of the shape matrix $\mathbf{Q}$. The minimum volume ellipsoid containing the $\{x^{(p)}\}_{p=1}^m$

can be computed by the following nonconvex formulation,

$$\begin{aligned} \min_{\mathbf{Q},\bar{x}} \ & \det(\mathbf{Q})^{-\frac{1}{2}} \\ \text{s.t.} \ & (x^{(p)} - \bar{x})^\intercal \mathbf{Q}(x^{(p)} - \bar{x}) \leq 1, \qquad p = 1, ..., m \\ & \mathbf{Q} \succ 0. \end{aligned} \tag{5.2}$$

For simplicity, assume the centered case where $\bar{x}$ is fixed to $\bar{x} = 0$. This is a superficial restriction as the uncentered problem can be exactly recovered by solving for a centered minimum volume enclosing ellipse in a lifted space [114, Section 2.3]. Hence we may generally consider problems of the centered form,

$$\begin{aligned} \min_{\mathbf{Q}} \ & -\operatorname{logdet}(\mathbf{Q}) \\ \text{s.t.} \ & (x^{(p)})^\intercal \mathbf{Q} x^{(p)} \leq 1, \qquad p = 1, ..., m \\ & \mathbf{Q} \succ 0. \end{aligned} \tag{5.3}$$

Several specialized algorithms exist for efficiently solving Equation (5.3). For example, Sun and Freund [112] propose an interior-point and active-set method that accurately solves moderate to large problems in a manner of seconds. In the monograph by Todd [114, Chapter 3], two first order methods – the Fedorov-Wynn-Frank-Wolfe algorithm and the Wolfe-Atwood algorithm – are detailed and analyzed along with modifications that scale to even larger dimension. In the numerical example in Section 5.5, we use the implementation of these algorithms provided in [114, Appendix B].

Datasets and ellipsoids

The inclusion of ellipsoid constraints for QOIs that require alternate feasibility criteria is straightforward. As an illustration, consider a QOI that is better fit or characterized by the tools of Bayesian history matching, i.e., stochastic emulation, implausibility, etc. Note that directly including the implausibility constraint $I_i(x) < 3$ into the dataset would lead to a challenging consistency analysis where provable guarantees may not be achievable. To circumvent this, the information in such constraints can be incorporated in an approximate fashion by generating samples in the non-implausible region as is typically done in history matching (Section 3.2) and then computing the minimum volume ellipsoid containing those retained samples. As this step only involves evaluating an emulator, a large number of samples may be used to characterize the non-implausible region. The ellipsoid can then be included in the dataset in lieu of the corresponding QOI's model-data constraint. In the same way, ellipsoids can be generated for other sample-based feasibility criteria, for instance, the regular B2BDC criteria with more sophisticated surrogate models, and incorporated into the dataset.

Let $\mathcal{J} \subseteq \{1, ..., N\}$ contain the indices of the QOIs which can be accurately represented through quadratic surrogate models and experimental bounds, i.e., a typical B2BDC setup.

The complement $\mathcal{J}^c = \{1, ..., N\} \backslash \mathcal{J}$ denotes those QOIs which are characterized through ellipsoids. The tools in Chapter 2 require little modification to accommodate the new constraints. As an example, the consistency measure in Equation (2.4) can be extended to,

$$\begin{aligned} C_{\mathrm{D}} := \max_{\gamma,\ x \in \mathcal{H}} \quad & \gamma \\ \text{s.t.} \quad & L_i + \frac{(U_i - L_i)}{2}\gamma \leq M_i(x_{\mathcal{A}_i}) \leq U_i - \frac{(U_i - L_i)}{2}\gamma, \quad \text{for all } i \in \mathcal{J} \\ & (x_{\mathcal{A}_k} - \bar{x}_{\mathcal{A}_k})^{\intercal} \mathbf{Q}_k (x_{\mathcal{A}_k} - \bar{x}_{\mathcal{A}_k}) \leq 1 - \gamma, \quad \text{for all } k \in \mathcal{J}^c, \end{aligned} \tag{5.4}$$

and the solution techniques described in Section 2.7 are unchanged. Note that the ellipsoids are only constructed in the active variables of the corresponding QOI. Hence, when viewed in the full dimensional space, these constraints carve out elliptic cylinders (the inactive dimensions are constrained by only the prior $\mathcal{H}$). As before, if the consistency measure is negative, then we have a proof of inconsistency – there is no common location in the parameter space that satisfies the prior, the model-data constraints, and the ellipsoid constraints. The optimizations required for prediction and other inferences can be reformulated in a similar manner.

Interpretation of consistency with ellipsoidal approximations

John's Theorem [61, Theorem III] states that for a given collection of points $\{x^{(p)}\}_{p=1}^{m}$, the minimum volume enclosing ellipsoid $\mathcal{E}$ satisfies the following chain of containments [114, Theorem 2.13]:

$$\bar{x} + \frac{1}{n}(\mathcal{E} - \bar{x}) \subseteq \mathrm{conv}(\{x_1, ..., x_m\}) \subseteq \mathcal{E} \tag{5.5}$$

where $\mathrm{conv}(\cdot)$ refers to the convex hull of the argument and

$$\bar{x} + \frac{1}{n}(\mathcal{E} - \bar{x}) = \{x : (x - \bar{x})^{\intercal} Q (x - \bar{x}) \leq \frac{1}{n^2}\} \tag{5.6}$$

is a scaling relative to the center point.

The second containment in Equation (5.5) is trivial, but when combined with the first, highlights the approximate nature of the minimum volume enclosing ellipsoid. In the context of B2BDC, let us temporarily fix a QOI i. Let $\{x_{\mathcal{A}_i}^{(p)}\}_{p=1}^{m_i}$ refer to the feasible samples used to generate the minimum volume ellipsoid $\mathcal{E}_i$. Additionally, define $n_i = |\mathcal{A}_i|$ where $|\cdot|$ refers to cardinality. In the subspace defined by the active variables, Equation (5.5) holds with the appropriate adjustments, e.g., n_i replaces n, etc. In the full dimensional space, i.e., $\mathbb{R}^n$, a similar containment inequality holds, albeit with cylinders – whose cross-sections are given by $\bar{x}_{\mathcal{A}_i} + \frac{1}{n_i}(\mathcal{E}_i - \bar{x}_{\mathcal{A}_i})$, $\mathrm{conv}(\{x_{\mathcal{A}_i}^{(1)}, ..., x_{\mathcal{A}_i}^{(m_i)}\})$, and $\mathcal{E}_i$ – replacing their counterparts in the inequality. Hence taking the intersection of either the minimum volume ellipsoids (as in Equation (5.4)) or their scaled versions provides information regarding the intersection of the convex hulls of the feasible samples. Indeed, a diagnosis of inconsistency with the

minimum volume ellipsoids proves the hulls do not intersect. A diagnosis of consistency with the scaled ellipsoids demonstrates that the hulls overlap.

It is readily apparent that the quality of a minimum volume bounding ellipsoid $\mathcal{E}_i$ depends on two aspects: how well the feasible samples $\{x_{\mathcal{A}_i}^{(p)}\}_{p=1}^{m_i}$ characterize $\mathcal{F}_i$ and how well $\mathcal{E}_i$ characterizes $\text{conv}(\mathcal{F}_i)$. Note, since $\{x_{\mathcal{A}_i}^{(p)}\}_{p=1}^{m_i} \subseteq \mathcal{F}_i$, we have that $\text{conv}(\{x_{\mathcal{A}_i}^{(1)}, ..., x_{\mathcal{A}_i}^{(m_i)}\}) \subseteq \text{conv}(\mathcal{F}_i)$. There is always the risk that despite $\mathcal{E}_i$ overapproximating the convex hull of the feasible samples as in Equation (5.5), it may still truncate an unsampled region of $\mathcal{F}_i$. The utility of this strategy relies heavily on quality of the feasible samples.

5.4 Generating additional quadratic constraints using support vector machines

A common challenge when using a bounding geometry is that it is often an overapproximation to the region being studied, as described in Section 5.3. In the context of B2BDC, this implies the actual feasibility constraint may be only weakly implemented using the ellipsoidal approximation. This difficulty is exacerbated as the problem dimension increases. For example, Zamora-Sillero et al. [124] note that the volume ratio of a nonconvex set to the associated bounding ellipsoid is usually very small in high dimensional settings. In previous B2BDC work by Russi et al. [99], this was demonstrated in the context of prediction. The authors computed various bounding geometries to the entire feasible set and observed significant degradation in the quality of predictions when those bounding regions were used instead of the feasible region. Despite this, incorporating a single or multiple ellipsoids representing individual or grouped QOIs as additional constraints (on top of $\mathcal{H}$) can still promote a significant reduction in the uncertain parameter space.

There are multiple strategies for alleviating this overapproximation problem. One solution discussed earlier is to implement B2BDC in an iterative fashion. This relies on the typical B2BDC surrogate models accurately fitting QOIs that were initially represented through ellipsoids over the reduced parameter region. Another solution is to include additional constraints into the dataset to eradicate as much excess as possible in the overapproximation. Naturally, these two approaches can be combined. In order to maintain the existing computational infrastructure, the aim is to formulate these additional constraints as convex or nonconvex quadratic inequalities. This section serves as a brief aside to discuss how support vector machines (SVMs) [26, 20, 88] can be used to generate additional quadratic constraints.

Let $\{(x^{(p)}, y^{(p)})\}_{p=1}^{m}$ be a collection of labeled training data with class labels $y^{(p)} \in \{-1, +1\}$. As a convention, label feasible points with $+1$ and infeasible points with -1. The dual form of kernel SVM classification can be expressed by the following quadratic

program [20, 88]:

$$\begin{aligned} \max_{\alpha_i} \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \\ \text{s.t. } 0 \leq \alpha_i \leq C \\ \sum_i \alpha_i y_i = 0 \\ i, j = 1, \ldots, m \end{aligned} \tag{5.7}$$

where $K(\cdot, \cdot)$ is an appropriate kernel function. In what follows the kernel will correspond to a polynomial kernel of degree 2, i.e., $K(x, z) = (1 + x^\intercal z)^d$ with $d = 2$. The optimal decision function that classifies inputs can be constructed after solving Equation (5.7) by $\text{class}(x) = \text{sign}(y(x))$ where

$$y(x) = \sum_{\{i:\alpha_i > 0\}} \alpha_i y_i K(x_i, x) + b. \tag{5.8}$$

The sum in the above expression is taken over all support vectors ($\{x_i : \alpha_i > 0\}$) and the offset b is given by picking any $i^\star$ such that $0 < \alpha_{i^\star} < C$ and solving $y(x_{i^\star}) = 1$. For quadratic kernels, the decision boundary can be converted to a more conventional quadratic form. Without loss of generality, assume that the data indices $i = 1, \ldots, N_s$ corresponds to the support vectors. Then,

$$\begin{aligned} \sum_{\{i:\alpha_i > 0\}} \alpha_i y_i K(x_i, x) + b &= \sum_{i=1}^{N_s} \alpha_i y_i (1 + x_i^\intercal x)^2 + b \\ &= \begin{bmatrix} 1 + x_1^\intercal x \\ \vdots \\ 1 + x_{N_s}^\intercal x \end{bmatrix}^\intercal \begin{bmatrix} \alpha_1 y_1 & & \\ & \ddots & \\ & & \alpha_{N_s} y_{N_s} \end{bmatrix} \begin{bmatrix} 1 + x_1^\intercal x \\ \vdots \\ 1 + x_{N_s}^\intercal x \end{bmatrix} + b \\ &= \begin{bmatrix} 1 \\ x \end{bmatrix}^\intercal \begin{bmatrix} 1 & \ldots & 1 \\ x_1 & \ldots & x_{N_s} \end{bmatrix} \begin{bmatrix} \alpha_1 y_1 & & \\ & \ddots & \\ & & \alpha_{N_s} y_{N_s} \end{bmatrix} \begin{bmatrix} 1 & \ldots & 1 \\ x_1 & \ldots & x_{N_s} \end{bmatrix}^\intercal \begin{bmatrix} 1 \\ x \end{bmatrix} + b \\ &= \begin{bmatrix} 1 \\ x \end{bmatrix}^\intercal W \begin{bmatrix} 1 \\ x \end{bmatrix} \end{aligned} \tag{5.9}$$

where W in the last equality is defined by the corresponding middle matrix in the previous step with the offset accounted for in the constant term. Hence, the inequality

$$y(x) = \begin{bmatrix} 1 \\ x \end{bmatrix}^\intercal W \begin{bmatrix} 1 \\ x \end{bmatrix} \geq 0 \tag{5.10}$$

characterizes the region with feasible labels and may be directly incorporated into the dataset as an additional quadratic constraint.

5.5 Example: Chemical Spill Model

The chemical spill example described in [15, 113] illustrates a scenario in which a chemical pollutant is released at two locations in an infinitely long one-dimensional channel. The concentration C of the chemical in the channel is governed by:

$$C(s,t;M,D,L,\tau) = \frac{M}{\sqrt{4\pi Dt}} \exp\left[\frac{-s^2}{4Dt}\right] + \frac{M}{\sqrt{4\pi D(t-\tau)}} \exp\left[\frac{-(s-L)^2}{4D(t-\tau)}\right] \mathbb{1}_{\{\tau<t\}}(t), \quad (5.11)$$

where (s,t) refers to the spatial and temporal dimensions of the problem. The essential setup is that a spill of mass M occurs in the channel at $(s,t) = (0,0)$. This spill is captured by the first term in Equation (5.11). At some later time τ another spill of the same mass occurs at a new location L. This second spill is modeled by the second term, which only becomes active after time τ. The parameter D specifies the diffusion rate associated with the channel. The parameter vector $x^\intercal = \begin{bmatrix} M & D & L & \tau \end{bmatrix}^\intercal$ contains all of the uncertain parameters, with prior bounds given by a hyperrectangle $\mathcal{H}$: $M \in [7,13]$, $D \in [0.2, 0.12]$, $L \in [0.01, 3]$, and $\tau \in [30.01, 30.295]$. In this section, we illustrate how minimum volume ellipsoids can be used to initiate a waves procedure.

Experimental measurements were simulated by evaluating Equation (5.11) at a series of $N = 75$ coordinates, denoted by $\{(s_i,t_i)\}_{i=1}^N$, with the parameters fixed at a hidden ground truth $x_\star^\intercal = \begin{bmatrix} 10 & 0.7 & 1.5 & 30.16 \end{bmatrix}^\intercal$. The nominal measurements for $C(s,t;x_\star)$ are shown below in Figure 5.1. To set the example up, we assume each measurement was assessed with $\pm 5\%$ uncertainty, producing experimental bounds of the form $L_i = 0.95 \times C(s_i,t_i;x_\star)$ and $U_i = 1.05 \times C(s_i,t_i;x_\star)$ for $i = 1, ..., N$.

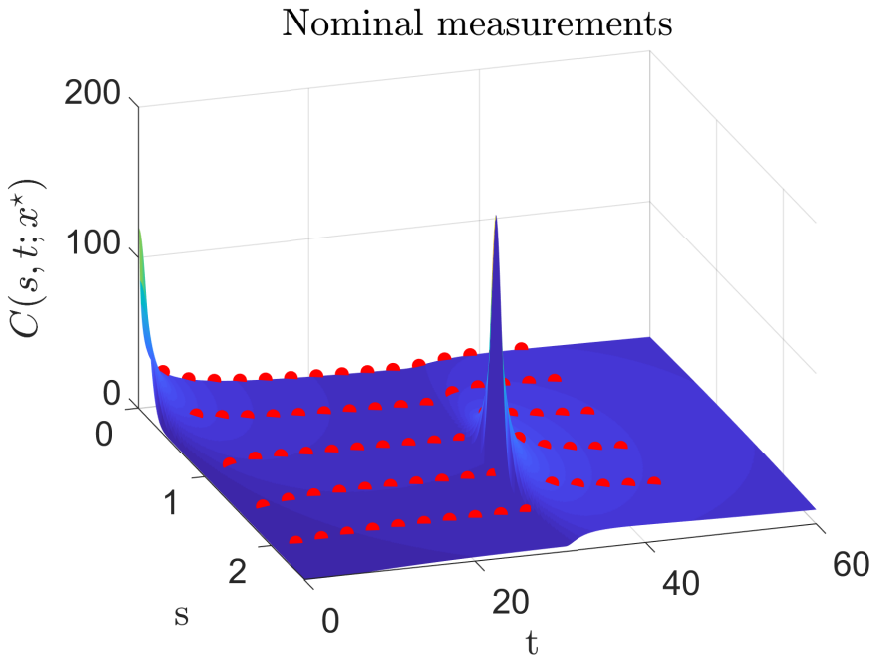

Figure 5.1: Nominal measurements of the QOIs of the ground truth system $C(s,t;x_\star)$ are displayed in red.

The first step in the analysis involved fitting surrogate models $M_i(x)$ to the simulation $C(s_i,t_i;x)$ over the initial prior parameter domain $\mathcal{H}$ for each QOI $i = 1,...,N$. This was accomplished by evaluating Equation (5.11) on a Latin hypercube (LH) design of 500 points over $\mathcal{H}$, partitioning the runs into both training (40%) and test (60%) cases, fitting the surrogate models on the training case, and then evaluating the errors on the combined set of runs. In order to gauge the quality of the fit, we computed the EEF as defined previously in Equation (3.6) with the cutoff set as $\alpha = 1.1$.

The outcome of this initial procedure for two types of surrogate models – quadratic and Gaussian process (GP) – is shown in Figure 5.2. The GP emulators were fit using the MATLAB [73] routine `fitrgp`. After this initial stage, none of the 75 QOIs were well approximated by quadratic fits. In contrast, several of GP emulators passed the EEF criteria. Note that in this setting, the typical waves procedure discussed in Chapter 3 would not be implementable.

To proceed with the analysis, we imposed the α-cutoff on the EEFs, which retained 35 GP surrogates. Following the approach described in Section 5.3, we then constructed a dataset comprised of the corresponding 35 minimum volume enclosing ellipsoids, one for each of the GP surrogates. The points used to fit the ith QOI's ellipsoid were generated by performing rejection using an initial design of 5×10^6 LH samples with the feasibility criteria

$$L_i - 2\sqrt{\mathrm{Var}(M_i(x))} \leq \mathbb{E}[M_i(x)] \leq U_i + 2\sqrt{\mathrm{Var}(M_i(x))}. \tag{5.12}$$

Note that in the above criteria, points that may be infeasible with respect to the mean evaluation are still labeled as feasible as long as the deviation can be explained by the emulator uncertainty. In the context of this example, the motivation for such criteria was to avoid eliminating points that may become useful in future iterations as the surrogate models improve. The resulting dataset (now made up of 35 ellipsoidal constraints) was consistent, as verified by computing the scalar consistency measure. Moreover, of the 5×10^6 LH points, roughly 10.1% were feasible with respect to the dataset, indicating a 89.9% reduction in volume from the initial prior region. Following the notation established in the preceding section, this updated feasible set produce in wave 1 is denoted as $\mathcal{F}^2$.

The procedure described above was repeated for additional waves $k = 2, 3, 4$, where the feasible region established at the completion of wave k, i.e., $\mathcal{F}^{k+1}$, served as the prior during the following wave $k+1$. For example, during the second wave ($k = 2$), surrogate models and error estimates were constructed using a set of 500 points from $\mathcal{F}^2$. In many cases, such as the example in Section 3.6, sophisticated sampling strategies are required to generate points in the feasible set. As the dimensionality of this example is comparatively small, samples were generated during each wave using the following procedure:

1. first, compute a bounding hyperrectangle to $\mathcal{F}^k$, e.g., as defined in Equation (2.6) and Section 2.8;

2. second, construct a LH design of 5×10^6 points in this region; and

3. third, retain only the samples that are feasible with respect to $\mathcal{F}^k$.

Of the retained samples, 500 were randomly selected and used for fitting surrogate models and computing error estimates. Those QOIs that passed the EEF cutoff were appended to the dataset, either as model-data constraints involving quadratic surrogate models or minimum volume ellipsoids.

Throughout the entire procedure, constraints in the dataset were continually accumulated. Maintaining this history ensures that at the conclusion of wave k the updated feasible set $\mathcal{F}^{k+1}$ is always confined to the prior region $\mathcal{F}^k$, i.e., $\mathcal{F}^{k+1} \subseteq \mathcal{F}^k$. Recall, $\mathcal{F}^k$ is exactly where the surrogates were both constructed and estimated as accurate. The surrogate model EEF criteria for waves 1 and 2 are shown below in Figure 5.2, and for waves 3 and 4 in Figure 5.3.

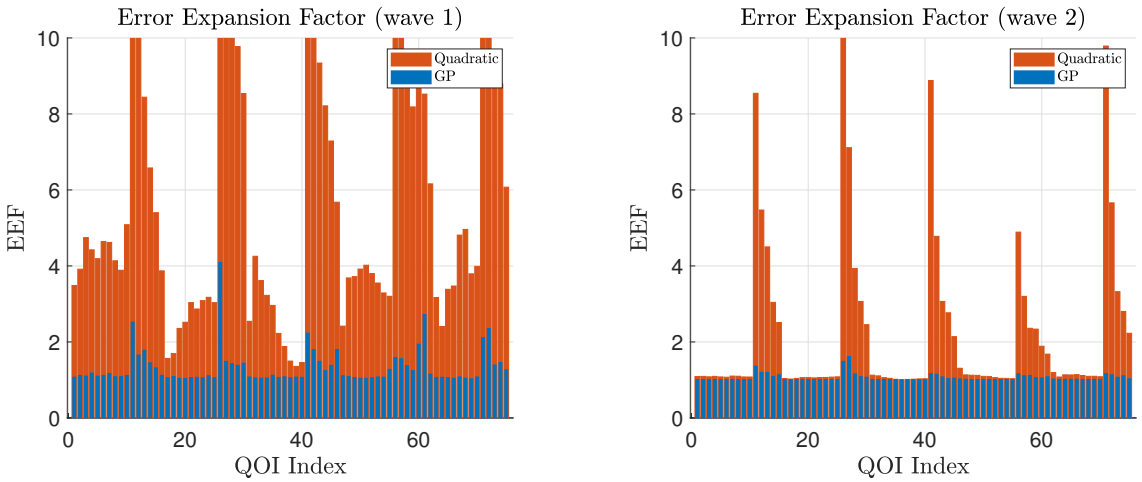

Figure 5.2: Left, EEFs for the first wave. Right, EEFs for the second wave.

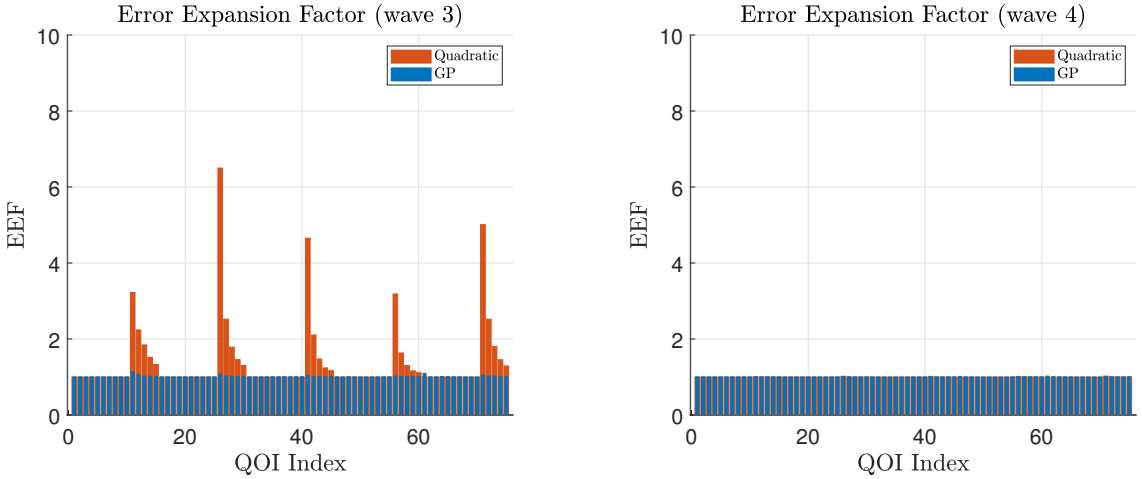

Figure 5.3: Left, EEFs for the third wave. Right, EEFs for the fourth wave.

As a general trend, the EEFs decreases for both classes of surrogate models as the waves progress. By the end of the fourth wave, all of the N QOIs are estimated as having accurate quadratic approximations. The composition of the dataset after each wave is compiled below in Table 5.1.

Table 5.1: Composition of the dataset after each wave. The number of ellipsoids corresponds to the number of GP surrogates that passed the error criteria during the kth wave. The volume fraction is calculated using a baseline collection of 5×10^6 LH samples from $\mathcal{H}$.

Wave	$k = 1$	$k = 2$	$k = 3$	$k = 4$
# quadratic surrogates added	0	37	50	75
# bounding ellipsoids added	35	23	24	0
# total QOIs represented	35	60	74	75
# total dataset constraints	35	95	169	244
$\text{vol}(\mathcal{F}^{k+1})/\text{vol}(\mathcal{H})$	0.101	2.29×10^{-3}	1.50×10^{-4}	3.54×10^{-5}

Needless to say, the dataset was found to be consistent during each iteration, as verified by both computing a consistency measure and the presence of feasible samples. Moreover, the "true" parameter configure $x_\star$ is feasible with respect to the final dataset, i.e., $x_\star \in \mathcal{F}^5$. Importantly, the predictions with the final dataset demonstrate close agreement when examined using the underlying computational model, i.e., Equation (5.11). Model predictions for all QOIs with a random selection of 1000 feasible parameter vectors are shown below in Figure 5.4.

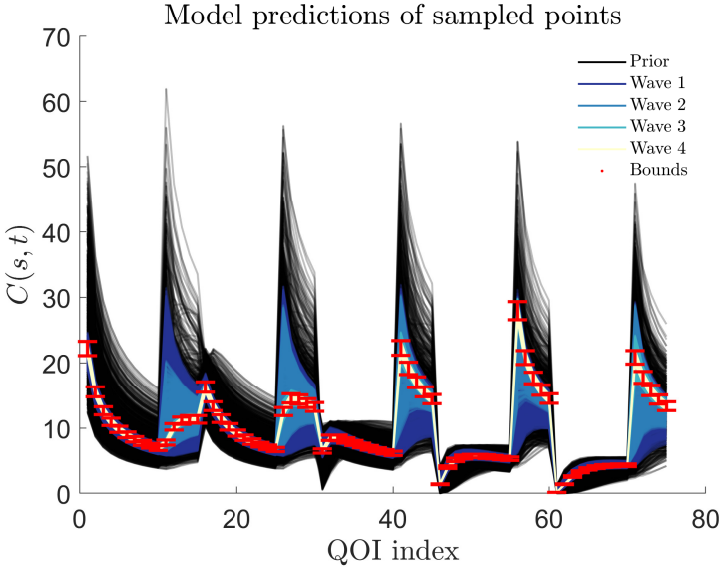

Figure 5.4: QOI traces of 1000 randomly selected feasible parameter vectors after the kth wave (i.e., samples from from $\mathcal{F}^{k+1}$). These traces were generated by directly evaluating Equation (5.11). The experimental bounds are shown in red.

The model-based traces that result from the 4th wave generally respect the experimental bounds. Of the 1000 samples displayed in Figure 5.4, 942 are feasible with respect to the underlying simulation. The remaining 58 samples, however, violate certain experimental bounds. For example, the feasible set contains 10 samples that under-predict the QOI 61's experimental bounds and 35 that under-predict QOI 71. This is an unavoidable consequence of using surrogate models.

The extent of the feasible set after the wave k dataset has been constructed can be investigated using different tools. For example, the posterior bounds on the parameters can be calculated by computing Equation (2.6) for each uncertain parameter $j = 1, ..., 4$. The posterior outer bounds for the final dataset are tabulated below in Table 5.2.

Table 5.2: Prior bounds and posterior outer bounds on the uncertain parameters at the conclusion of wave 4.

Parameters	Prior		Posterior	
	Lower	Upper	Lower	Upper
M	7	13	9.5	10.5
D	0.02	0.12	0.0685	0.0715
L	0.01	3	1.45	1.57
τ	30.01	30.295	30.01	30.295

The posterior bounds for the parameters M, D and L are reduced compared to the prior. The bounds for τ, however, show no improvement. Interestingly, of 5×10^6 LH samples generated in the posterior hyperrectangle, approximately 18.5% were feasible. This highlights the fact that the feasible set has structure unaccounted for by just the posterior intervals. Pairwise correlations among the variables are shown below in Figure 5.5.

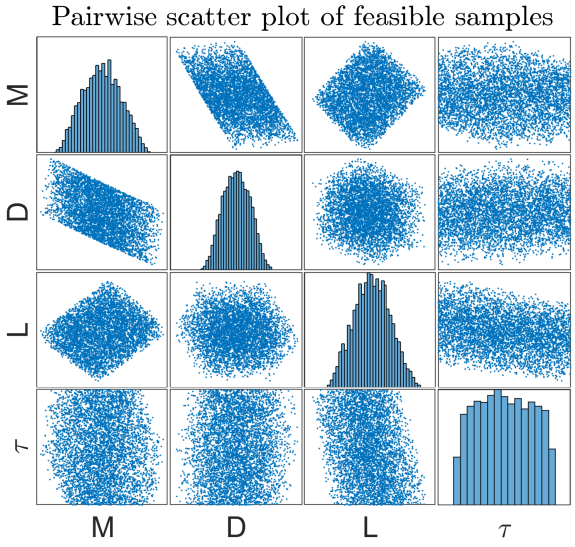

Figure 5.5: Pairwise scatter plot of 5000 randomly selected feasible parameter vectors. Note that the axes have been set to the corresponding posterior outer bounds.

5.6 Chapter summary

Including alternate feasibility information through extra quadratic constraints is a natural addition to the B2BDC framework. In this chapter, we investigated the inclusion of minimum volume bounding ellipsoids as an encapsulation of point-wise feasibility criteria. By example, it was demonstrated that such an approach can be combined with the iteration-based strategy of Chapter 3. This merging of techniques not only retains the existing computational infrastructure, but also extends the applicability of iteration-based B2BDC strategies.

Chapter 6

Sensitivities and the Scalar Consistency Measure

6.1 Introduction

The previous chapters develop iteration-based strategies for constructing a B2BDC dataset. As demonstrated, these techniques extend the domain of the methodology and open doors to new applications. Within this iterative context, the notion of consistency was of vital importance. Inconsistency served as a terminal condition for the analysis – a provable demonstration that no feasible set exists and that further refinement is impossible. Moreover, inconsistency signals that the dataset is unsuitable for prediction. In such cases, the focus of any scientific endeavor must then shift inwards to analyzing the collective content of the dataset. The remainder of this thesis develops tools to aid in this essential task. A diagnosis of inconsistency establishes that the models, prior information, and experimental data are at odds. The questions of how to identify the sources of such conflict and how to resolve them presents a challenge to the scientific community. Previous work by Feeley et al. [37] demonstrated that the multipliers in the Lagrangian dual formulation of the SCM can be used to identify possibly flawed experimental data responsible for an inconsistency. Essentially, these multipliers are interpreted as sensitivity estimates to perturbations of the constraint bounds. In a study by Russi et al. [98], the use of sensitivities was extended to assessing the impact of prior and experimental bounds on prediction bounds. The basic ethos is this: if the conclusion of a consistency or prediction analysis is highly sensitive to the tentatively-entertained uncertainty in particular dataset constraints, then further investigation of the identified items is warranted. Without the intervention of a domain expert, however, it cannot be discerned whether the flagged bounds are: (a) truly problematic, e.g., incorrect experimental/prior data, or (b) highly influential, e.g., the result of a high quality and informative experiment. The outcomes of such an analysis may include a recommendation of future experiments to repeat or conduct.

This chapter serves two purposes. The first is a continued examination of inconsistency

with respect to the SCM. This material motivates the developments in Chapter 7. The second purpose is to expand on the computational machinery discussed earlier in Section 2.7 for NQCQPs. Much of the content in this chapter is an extension of prior work by Feeley et al. [37] and is partially documented in our recent publication [54].

6.2 Running examples

Throughout this chapter, we illustrate sensitivity-based strategies for resolving inconsistency using two example datasets from the field of combustion chemistry – GRI-Mech 3.0 [107] and the more recent DLR-SynG [105]. As will be shown, the GRI-Mech 3.0 dataset illustrates a successful application of these multiplier-based sensitivities. In contrast, the second dataset – DLR-SynG – demonstrates a pitfall with this approach. The successes and failures of the SCM-based method for resolving inconsistency, as displayed by these examples, motivates the development of an additional tool, which is introduced and examined in Chapter 7. The two running examples are summarized below and will be revisited throughout this and the following chapter.

GRI-Mech 3.0

The GRI-Mech 3.0 dataset was constructed to calibrate an underlying kinetic reaction model for pollutant formation in natural gas combustion [107]. The kinetic model is a system of ordinary differential equations (ODE) representing 325 reversible reactions among 53 different chemical species. The dataset consists of $N = 77$ QOIs and $n = 102$ model parameters. Each of the 77 quadratic surrogate models is developed in approximately 10 to 15 active variables. The application of B2BDC to this dataset has been extensively investigated in previous works [42, 37, 123, 41].

DLR-SynG

The second example, the DLR-SynG dataset, was constructed to analyze a predictive kinetic model for syngas combustion [105]. Similar to GRI-Mech 3.0, DLR-SynG is formulated as an ODE system representing 73 reactions in 17 different chemical species. The dataset is comprised of 159 model-data constraints in 55 uncertain parameters. The QOI surrogate models are also quadratic and developed in a subset of active variables. The experimental bounds were assessed by the authors and based on data reported in the literature [105].

6.3 Sensitivity through perturbation

As described in Section 2.7, the SDP and Lagrangian approaches form dual perspectives on developing convex relaxations to NQCQPs. In the context of the SCM, the multipliers of the Lagrangian relaxation can be interpreted as estimating the impact of perturbations

to particular constraint bounds. This is detailed below for the case in which the prior is postulated as a hyperrectangle $\mathcal{H} = [l_1, u_1] \times \ldots \times [l_n, u_n]$. For convenience, the following notation is used throughout this section:

$$h_j^u(x) = x_j - u_j, \quad h_j^l(x) = -x_j + l_j \qquad \text{for } j = 1, \ldots, n \tag{6.1}$$

$$h_i^U(x, \gamma) = M_i(x_{\mathcal{A}_i}) - U_i + \frac{U_i - L_i}{2}\gamma \qquad \text{for } i = 1, \ldots, N \tag{6.2}$$

$$h_i^L(x, \gamma) = -M_i(x_{\mathcal{A}_i}) + L_i + \frac{U_i - L_i}{2}\gamma \tag{6.3}$$

Let $\rho = \begin{bmatrix}(\rho^u)^\intercal & (\rho^l)^\intercal & (\rho^U)^\intercal & (\rho^L)^\intercal\end{bmatrix}^\intercal \in \mathbb{R}^{2n+2N}$ denote a vector of perturbations. Consider the perturbed version of the SCM shown below.

$$\begin{aligned} C_{\mathrm{D}}(\rho) := \max_{\gamma, x} \quad & \gamma \\ \text{s.t.} \quad & h_j^u(x) \le \rho_j^u, \qquad h_j^l(x) \le \rho_j^l, \quad \text{for } j = 1, \ldots, n. \\ & h_i^U(x, \gamma) \le \rho_i^U, \quad h_i^L(x, \gamma) \le \rho_i^L, \quad \text{for } i = 1, \ldots, N. \end{aligned} \tag{6.4}$$

Setting $\rho > 0$ relaxes the associated prior/experimental bounds, e.g., $l_i - \rho_j^l \le x_i \le u_i + \rho_j^u$, whereas $\rho < 0$ provides a tightening. Naturally $C_D(0)$ recovers the original SCM in Equation (2.4). Following similar steps as in the Lagrangian dual relaxation (Section 2.7) while accounting for the maximization leads to the following upper bound:

$$\begin{aligned} C_{\mathrm{D}}(\rho) \le \overline{C}_{\mathrm{D}}(\rho) := \min_{\lambda \ge 0} \max_{\gamma, x} \gamma &+ \sum_{j=1}^{n} (\lambda_j^u(\rho_j^u - h_j^u(x)) + \lambda_j^l(\rho_j^l - h_j^l(x))) \\ &+ \sum_{i=1}^{N} (\lambda_i^U(\rho_i^U - h_i^U(x, \gamma)) + \lambda_i^L(\rho_i^L - h_i^L(x, \gamma))). \end{aligned} \tag{6.5}$$

As discussed in earlier sections, the computation of $\overline{C}_{\mathrm{D}}(\rho)$ can be cast as an SDP if quadratic surrogate models are employed. Let $\lambda^\#$ be the minimizer of the Lagrangian relaxation to the SCM, i.e., the minimizer found by computing $\overline{C}_{\mathrm{D}}(0)$. Then, $\lambda^\#$ is suboptimal with respect to Equation (6.5). Hence for all ρ, the following holds,

$$\begin{aligned} C_{\mathrm{D}}(\rho) &\le \overline{C}_{\mathrm{D}}(\rho) \\ &\le \max_{\gamma, x} \gamma + \sum_{j=1}^{n} ((\lambda_j^u)^\#(\rho_j^u - h_j^u(x)) + (\lambda_j^l)^\#(\rho_j^l - h_j^l(x))) \\ &\qquad + \sum_{i=1}^{N} ((\lambda_i^U)^\#(\rho_i^U - h_i^U(x, \gamma)) + (\lambda_i^L)^\#(\rho_i^L - h_i^L(x, \gamma))) \\ &= \overline{C}_{\mathrm{D}}(0) + \sum_{j=1}^{n} \left((\lambda_j^u)^\#\rho_j^u + (\lambda_j^l)^\#\rho_j^l\right) + \sum_{i=1}^{N} \left((\lambda_i^U)^\#\rho_i^U + (\lambda_i^L)^\#\rho_i^L\right). \end{aligned} \tag{6.6}$$

Equation (6.6) demonstrates that the perturbed SCM is bounded above by an expression affine in the perturbations with intercept $\overline{C}_{\mathrm{D}}(0)$ and slopes given by $\lambda^{\#}$.

The interpretation of Equation (6.6) is most efficiently understood when considering a provably inconsistent dataset with $C_{\mathrm{D}} < 0$ ($\overline{C}_{\mathrm{D}} < 0$ is a sufficient condition). Suppose $\lambda_k^{\#} = 0$ for some index k, which may refer to either a prior bound or an experimental bound. Then the dataset remains inconsistent ($C_D(\rho) < 0$) for arbitrarily large perturbations to bound k, meaning the consistency is insensitive to the kth constraint's uncertainty. In general, large multipliers $\lambda_k^{\#}$ indicate that the SCM is highly responsive to corresponding constraint tightenings ($\rho_k < 0$) and small $\lambda_k^{\#}$ indicate that the SCM is rather unresponsive to corresponding constraint relaxations ($\rho_k > 0$).

A challenge with the above setup is that the Lagrange multipliers are not directly comparable among different QOIs and parameters due to differences in units. To address this, the sensitivities are often represented by the scaled multipliers $s_k = \lambda_k^{\#}(u_k - l_k)$, where $(u_k - l_k)$ is the corresponding uncertainty width. Equation (6.6) may be rewritten in terms of these new sensitivities,

$$C_{\mathrm{D}}(\rho) \leq \overline{C}_{\mathrm{D}}(0) + \sum_{j=1}^{n} \left(s_j^u \frac{\rho_j^u}{u_j - l_j} + s_j^l \frac{\rho_j^l}{u_j - l_j} \right) + \sum_{i=1}^{N} \left(s_i^U \frac{\rho_i^U}{U_i - L_i} + s_i^L \frac{\rho_i^L}{U_i - L_i} \right). \tag{6.7}$$

The sensitivities s_k are thus coefficients to relative perturbations (as opposed to absolute) in the corresponding constraint's uncertainty interval.

Computing the dual $\overline{C}_{\mathrm{D}}$ and its associated Lagrange multipliers can often be accomplished using the tools of convex optimization [19]. For the surrogate models described in Section 2.7, the dual computation can be addressed using SDP techniques. Additional aspects related to computations and prediction are detailed in prior work [37, 35, 98].

6.4 A strategy for resolving inconsistency

Reliable scientific conclusions should be robust relative to small perturbations of the prior and experimental data. Thus, roughly speaking, we should not expect the SCM $C_{\mathrm{D}}(\rho)$ to change significantly for small ρ. In the case of an inconsistent dataset, the sensitivities described above can be used to flag possibly problematic model-data constraints as suggested by Feeley et al. [37].

The use of SCM-sensitivities to flag constraints can be cast into a general strategy for resolving inconsistent datasets in B2BDC. This approach is iterative and summarized in Algorithm 1. Note that since the SCM is formulated as a maximization, the terminal condition for the iteration is based on a local maxima for the SCM – i.e., a verification of consistency.

Algorithm 1 Resolving inconsistency through sensitivity

1: **Input:** Inconsistent dataset D with $C_D \in [\underline{C}_D, \overline{C}_D]$ and $\overline{C}_D < 0$.
2: **while** $\underline{C}_D < 0$ **do**
3: Extract sensitivities from $\overline{C}_D$ computation.
4: Define a new dataset $\tilde{D}$ by removing "high sensitivity" model-data constraints.
5: Update $D \leftarrow \tilde{D}$.
6: Recompute $[\underline{C}_D, \overline{C}_D]$.
7: **end while**
8: **Output:** Consistent dataset D and the removed constraints.

We note that some of the steps in Algorithm 1 are ambiguous. For example, how to differentiate between "high" and "low" sensitivity is unspecified. A simple approach is to remove the constraint with the highest sensitivity at each iteration of the algorithm. Alternatively, one could remove the top two, top three, etc.

6.5 Strengthening the relaxation via redundant constraints

The determination of consistency is based on computing two quantities which bracket the actual solution $\underline{C}_D \leq C_D \leq \overline{C}_D$. Therefore, the consistency of a dataset can be verified by demonstrating $\underline{C}_D \geq 0$. In contrast, inconsistency is proven by showing $\overline{C}_D < 0$. In cases where $\underline{C}_D < 0 \leq \overline{C}_D$, no conclusion about consistency can be reached. Recall that this indeterminacy is not a characteristic of the dataset, but rather due to the solution techniques used to approximate the SCM.

The strategy for computing $\underline{C}_D$ described in Section 6.3 (and Section 2.7) may not be enough to draw a conclusion regarding consistency. We illustrate this with the GRI-Mech 3.0 dataset. Here, the computation of $\overline{C}_\mathrm{D}$ is performed using CVX [52, 53], a package for formulating and solving convex programs, with the solver Sedumi [111]. The local solution $\underline{C}_\mathrm{D}$ is found by initializing the optimization routine `fmincon` in MATLAB [73] at 50 random points, in accordance with Section 2.7. This strategy brackets the SCM with $C_D \in [-0.37, 1]$. This result is indeterminate.

To address this challenge, Feeley et al. [37, 35] strengthen the formulation for $\overline{C}_D$ by incorporating additional quadratic constraints derived from the box constraints defining $\mathcal{H}$. This recycling of problem data produces seemingly redundant constraints which, after relaxation, lead to a tightened $\overline{C}_D$. The general idea of deriving valid inequalities from existing constraints is commonly applied in strengthening convex relaxations to nonconvex problems [104, 92, 4, 6], and more generally characterized by Positivstellensatz refutations [90, 68]. The impact of such redundant constraints is illustrated in the following paragraph.

Consider prior bounds on the uncertain parameters defined by the box constraints $l_j \leq$

$x_j \leq u_j$ for $j = 1, ..., n$. In Section 6.3, these constraints were rewritten as a pair of linear inequalities,

$$\begin{aligned} & x_j - u_j \leq 0 \\ & l_j - x_j \leq 0 \\ & \text{for } j = 1, ..., n, \end{aligned} \tag{6.8}$$

and directly incorporated into the Lagrangian relaxation. These $2n$ constraints may also be equivalently represented by n quadratic inequalities,

$$\begin{aligned} 0 &\geq -(x_j - u_j)(l_j - x_j) \\ &= x_j^2 - (u_j + l_j)x_j + u_j l_j \\ &\text{for } j = 1, ..., n. \end{aligned} \tag{6.9}$$

These new quadratic constraints are redundant with respect to the original problem as they are derived from existing inequalities. Despite this, their inclusion can lead to a significant tightening of the convex relaxation $\overline{C}_D$. In the context of the GRI-Mech 3.0 dataset, the solution becomes $\overline{C}_D = -0.26$. The inconsistency of the dataset is now certified, with $C_D \in [-0.37, -0.26]$. The inclusion of these redundant constraints, however, occurs cost of more detailed sensitivity information regarding the parameter bounds. This aspect is explained by Feeley [35, Section 4.3], in which it is shown that the ability to isolate sensitivities pertaining to individual parameter bounds, i.e., λ_j^l and λ_j^u, is lost when the additional constraints are incorporated into the dataset. This does not significantly influence Algorithm 1, where the emphasis is on identifying possibly problematic experimental bounds rather than parameter bounds. Note, flagged parameter bounds require additional consideration as uncertain parameters cannot be so easily untangled from the dataset constraints. Additionally, expanding parameter bounds implies that all affected QOIs must have their surrogate models refit.

Note that additional valid inequalities can be derived from the cross terms as well,

$$\begin{aligned} & x_k - u_k \leq 0 \\ & l_j - x_j \leq 0 \\ & \text{for } j, k = 1, ..., n \text{ with } j \neq k. \end{aligned} \tag{6.10}$$

Similarly, these linear inequalities can be converted into seemingly redundant quadratics,

$$\begin{aligned} 0 &\geq -(x_k - u_k)(l_j - x_j) \\ &= x_k x_j - l_j x_k - u_k x_j + u_k l_j \\ &\text{for } j, k = 1, ..., n, \ k \neq j. \end{aligned} \tag{6.11}$$

In cases where n is large, including all of the extra constraints can become computationally unmanageable. This challenge has been recognized in recent literature. For example, Ashraphijou et al. [6] propose an iterative strategy in which inequalities that are binding at optimality are incorporated and the optimization is resolved. The "binding at optimality"

is most easily verified in the context of the SDP relaxation of the SCM, as formulated in Section 2.7. This is described in a more general setting below.

Consider the case in which the prior is defined by a set of linear inequalities, i.e.,

$$\mathcal{H} = \{x \in \mathcal{R}^n : \mathbf{A}^\intercal \begin{bmatrix} 1 \\ z \end{bmatrix} \leq 0; z^\intercal = \begin{bmatrix} x^\intercal & \gamma \end{bmatrix}\} \tag{6.12}$$

where $\mathbf{A} \in \mathbb{R}^{(1+n)\times m}$. Note that the box constraints described earlier correspond to a particular choice of $\mathbf{A}$. Then the collection of all redundant constraints can be expressed as,

$$\mathbf{A}^\intercal \begin{bmatrix} 1 \\ z \end{bmatrix} \leq 0 \quad \Rightarrow \quad \mathbf{A}^\intercal \begin{bmatrix} 1 \\ z \end{bmatrix} \begin{bmatrix} 1 \\ z \end{bmatrix}^\intercal \mathbf{A} \geq 0 \tag{6.13}$$

where the inequalities on both sides are applied element-wise. In the context of the SDP relaxation, the seemingly redundant inequalities become extra linear constraints on the matrix variable Z,

$$\begin{aligned} &\mathbf{A}^\intercal \mathbf{Z} \mathbf{A} \geq 0 \\ &\mathbf{Z}_{11} = 1 \\ &\mathbf{Z} \succeq 0. \end{aligned} \tag{6.14}$$

Returning to the "binding at optimality" question, let $\mathbf{Z}^\star$ denote the optimal solution to the SDP relaxation of the SCM (without the extra inequalities). The valid inequalities that would actually strengthen the solution are those that are violated by $\mathbf{Z}^\star$, i.e, those for which $(\mathbf{A}^\intercal \mathbf{Z}^\star \mathbf{A})_{jk} < 0$.

6.6 Resolving inconsistency in GRI-Mech 3.0

The sensitivity-based strategy described in Section 6.4 was first applied to the GRI-Mech 3.0 dataset [37]. In that work, the two identified experimental bounds – corresponding to two different QOIs – were confirmed incorrectly reported after consulting the experimenters. As an outcome, these specific bounds were revised with updated uncertainties. A more recent sensitivity analysis of the dataset by You et al. [123] found a single QOI, labeled f5, to be a possible source for the inconsistency. In this section, we repeat this analysis in order to contrast with that of the DLR-SynG dataset, which is discussed in the following section. In the analysis described below, the valid inequalities in Equation (6.9) have been incorporated into the dataset.

An initial application of the SCM to the GRI-Mech 3.0 dataset certifies that the dataset is inconsistent, with $C_D \in [-0.37, -0.26]$. The QOI bounds with the ten largest sensitivities, as defined in Section 6.3, are shown in Figure 6.1. Recall that these sensitivities estimate the influence of bound perturbations on the value of the SCM.

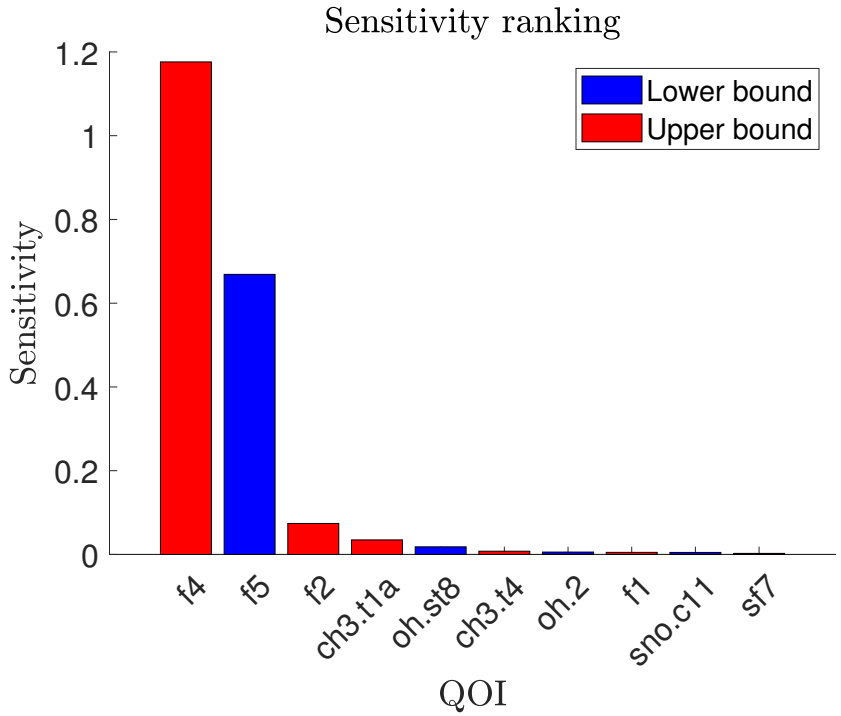

Figure 6.1: The SCM sensitivity ranking for GRI-Mech 3.0.

The top two QOIs – f4 and f5 – dominate the sensitivity ranking. Removing these flagged QOIs produces a provably consistent dataset. Moreover, a brief trial-and-error analysis reveals that removing just f5 results in consistency as well.

6.7 Resolving inconsistency in DLR-SynG

The situation for the DLR-SynG dataset is starkly different. An initial application of the SCM proves inconsistency, with $C_D \in [-1.90, -1.55]$. The corresponding sensitivities are shown in Figure 6.2.

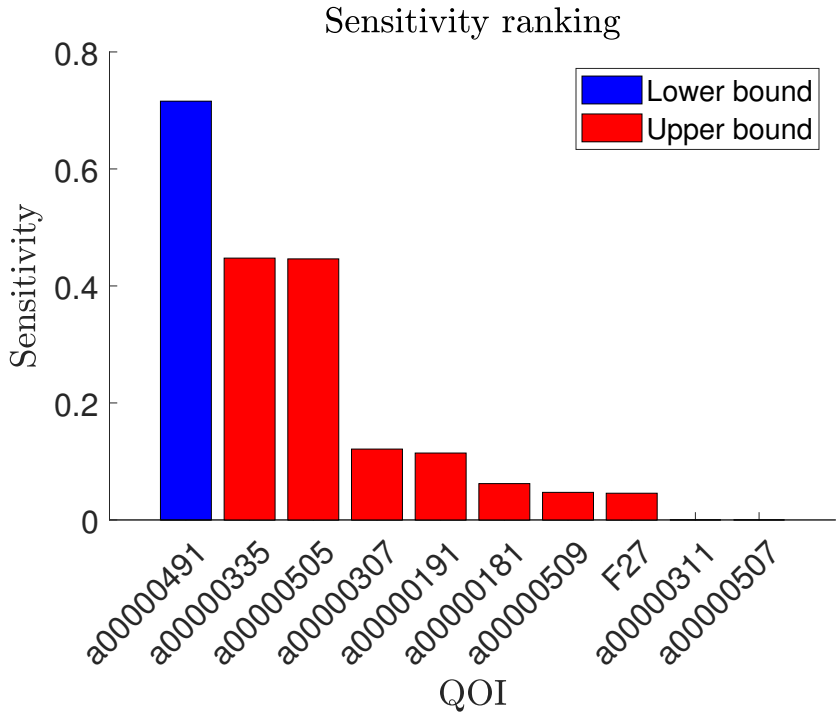

Figure 6.2: The SCM sensitivity ranking for DLR-SynG.

Initially the sensitivity ranking appears similar to that in the GRI-Mech 3.0 example. Further investigation reveals that removing any number of the flagged QOIs results in a dataset that is still provably inconsistent. For instance, deleting the QOI with the highest sensitivity and recomputing the SCM produces a new ranking, shown in Figure 6.3.

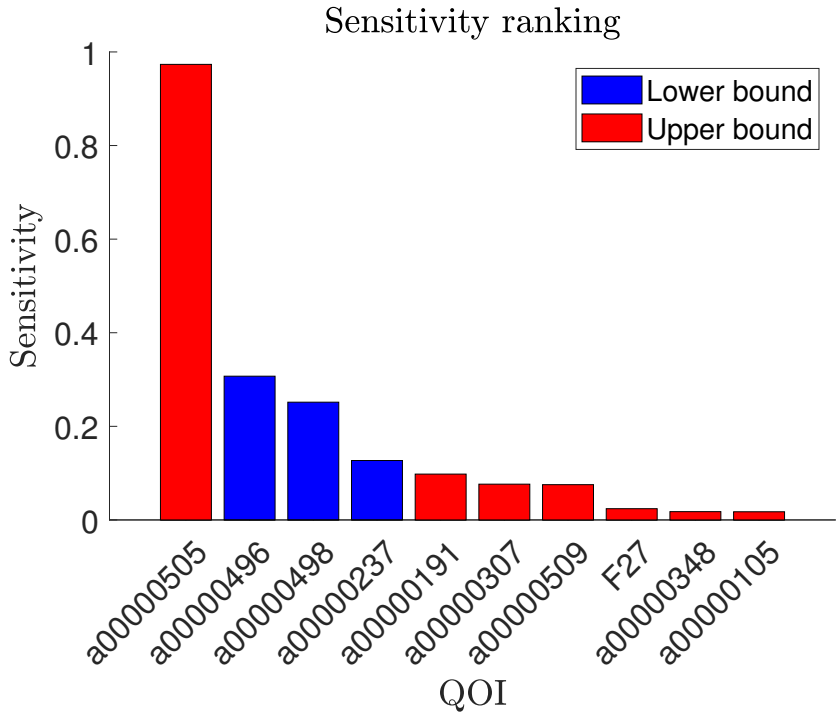

Figure 6.3: The SCM sensitivity ranking for DLR-SynG after the first constraint removal.

Note, many of the identified QOIs were also flagged during the first iteration. The ordering, however, has shifted with several new QOIs also appearing. For example, three new QOIs emerge with non-negligible sensitivities associated with the lower bound constraints. This process of calculating sensitivities and removing the QOI ranked highest was repeated until the terminal condition was met, i.e., the pruned subset of dataset QOIs was verified consistent. This procedure led to 56 model-data constraints being removed. It only takes a few iterations to notice that the scale of the inconsistency in DLR-SynG dwarfs that of GRI-Mech 3.0. After completing this entire procedure, one begins to recognize the ambiguity in Algorithm 1 – a natural question to ask is "what difference would it have made had I removed the second most sensitive constraint (as opposed to the top most sensitive constraint)?" Following this line of inquiry and deleting the second most sensitive constraint at each iteration led to a total of 54 constraints removed. At each stage, removing a different constraint may have resulted in different and potentially fewer future constraints removed. An alternative and extreme strategy would be to remove all constraints which appear in the top ten ranking. Doing so led to an iterative scheme which repeated seven times and resulted in the deletion of 70 model-data constraints, which certainly does not improve upon the previous results. With each analysis, the concern still remains: "could consistency be reached with fewer constraints removed?"

6.8 Chapter summary

The scalar consistency measure (SCM) is one of the pillars of the B2BDC framework. Historically, this tool has been used for both certifying consistency as well as diagnosing inconsistency. This latter task is addressed through computing the Lagrangian relaxation to the SCM, which produces multipliers that estimate the influence of perturbations to constraint bounds. These multipliers are referred to as SCM sensitivities. In this chapter, a formal sensitivity-based SCM procedure for handling dataset inconsistency was described. By example, we demonstrated that although the strategy can be successfully applied in some cases (GRI-Mech 3.0) there are other cases where the results are less than satisfying (DLR-SynG). The outcome of the DLR-SynG analysis motivates the developments presented in the next chapter.

Chapter 7

The Vector Consistency Measure

7.1 Introduction

The results of the previous chapter, particularly the questions posed in Section 6.7, motivate the need for new methods to handle dataset inconsistency. In this chapter, we propose a more refined tool, termed the *vector consistency measure* (VCM), for this purpose. This new approach can be considered an extension to the SCM and opens a new path in the B2BDC framework. Moreover, the introduction of weights into the VCM framework enables an even richer form of consistency analysis.

7.2 The Vector Consistency Measure

For an inconsistent dataset, the decision variable γ in the SCM acts as a single relaxation to all model-data constraints. For convenience, we reproduce the formulation in Equation (2.4) below,

$$\begin{aligned} C_{\mathrm{D}} := \max_{\gamma,\ x\in\mathcal{H}} \quad & \gamma \\ \text{s.t.} \quad & L_i + \frac{(U_i - L_i)}{2}\gamma \leq f_i(x) \leq U_i - \frac{(U_i - L_i)}{2}\gamma \\ & \text{for } i = 1, \ldots, N. \end{aligned} \tag{7.1}$$

This strategy can be made more flexible by instead allowing independent relaxations to each constraint – the idea being that not every bound actually needs adjustment in order to induce consistency. Hence, rather than having just a single scalar γ as a decision variable, multiple γ_i are in play. Under such circumstances, a natural question to ask of the inconsistency is the following: "of these relaxations, what is the fewest number required to reach consistency?"

This thought process leads to the optimization

$$
\begin{aligned}
V_{\|\cdot\|_0} := \min_{x,\Delta_L,\Delta_U,\delta_l,\delta_u} \; & \|\Delta_L\|_0 + \|\Delta_U\|_0 + \|\delta_l\|_0 + \|\delta_u\|_0 \\
\text{s.t.} \quad & L_i - \Delta_L^{(i)} \leq M_i(x) \leq U_i + \Delta_U^{(i)} \qquad \text{for } i = 1, ..., N \\
& l_j - \delta_l^{(j)} \leq x_j \leq u_j + \delta_u^{(j)} \qquad \text{for } j = 1, ..., n
\end{aligned} \tag{7.2}
$$

where the function $\|\cdot\|_0$, sometimes called the 0-norm, expresses the number of nonzero entries in the argument. The auxiliary variables Δ and δ have been introduced as relaxations to both model-data constraints and parameter bounds. If a dataset is consistent, then no relaxation is required and the optimal value is zero. For an inconsistent dataset, solution of the above problem provides three pieces of information: the smallest number of bound adjustments required to reach consistency, a collection of optimal relaxations, and a corresponding parameter vector that becomes consistent. Going further and actually removing the model-data constraints flagged with nonzero relaxations results in a pruned dataset which contains the largest subcollection of model-data constraints (from the original dataset) that are together consistent. In general, we note that the relaxation variables in Equation (7.2) can be taken as either strictly independent or supplemented with additional constraints to account for known or prescribed dependencies. For example, in certain cases one may wish to supply explicit correlations, like $\Delta_U^{(1)} \leq \Delta_U^{(2)}$, or restrictions, like $\Delta_U^{(1)} = 0$ and $\Delta_U^{(2)} \leq 0.01$.

The above optimization problem is difficult to solve as the 0-norm is nonconvex and even linear problems with this objective are NP-Hard [79]. To address the nonconvexity, we use the standard technique of replacing the 0-norm with the 1-norm, a well-known convex heuristic for sparsity [32, 22, 23]. This modification produces

$$
\begin{aligned}
V_{\|\cdot\|_1} := \min_{x,\Delta_L,\Delta_U,\delta_l,\delta_u} \; & \|\Delta_L\|_1 + \|\Delta_U\|_1 + \|\delta_l\|_1 + \|\delta_u\|_1 \\
\text{s.t.} \quad & L_i - \Delta_L^{(i)} \leq M_i(x) \leq U_i + \Delta_U^{(i)} \qquad \text{for } i = 1, ..., N \\
& l_j - \delta_l^{(j)} \leq x_j \leq u_j + \delta_u^{(j)} \qquad \text{for } j = 1, ..., n
\end{aligned} \tag{7.3}
$$

which we term the *vector consistency measure* (VCM) to differentiate from the SCM's scalar relaxation. Note that in moving to a 1-norm objective, the magnitudes of the relaxations are now penalized. In the context of a B2BDC analysis, we emphasize that the answer to the above problem should be interpreted as a theoretical reference point; whether relaxing the experimental bounds in the prescribed manner is justifiable or whether that relaxation signals a deficiency of the underlying physical model f is to be addressed by domain science. This aspect is emphasized in the following paragraph. What the analysis reveals is that directly implementing the relaxations, removing the model-data constraints with nonzero relaxations, or some combination of those actions are all strategies that result in a consistent dataset. This proposed tool aims at efficiently exploring the different ways an inconsistency can be resolved.

A consequence of the above formulation is that the VCM will never simultaneously relax both sides of an inequality constraint. As a result, there is no clear distinction between

a relaxation to an experimental bound and an additive bias to the corresponding model output. Without loss of generality, consider an inconsistent dataset with only model-data constraints. Let $\Delta_L^\star$ and $\Delta_U^\star$ be the VCM-optimal relaxations. Thus, the set

$$\mathcal{F}' = \{x : L_i - \Delta_L^{\star(i)} \leq M_i(x) \leq U_i + \Delta_U^{\star(i)},\ i = 1, ..., N\} \tag{7.4}$$

is nonempty. Suppose constraint k has nonzero relaxations associated with both bounds. Then there exists a $z \in \mathcal{F}'$ and such that $\Delta_L^{\star(k)} > 0$ and $\Delta_U^{\star(k)} > 0$. There are three cases to consider: $M_k(z) = L_k - \Delta_L^{\star(k)}$, $M_k(z) = U_k + \Delta_U^{\star(k)}$, or $L_k - \Delta_L^{\star(k)} < M_k(z) < U_e + \Delta_U^{\star(k)}$. In each case, a viable relaxation with smaller 1-norm can be found by either not including $\Delta_U^{\star(k)}$ (case 1), not including $\Delta_L^{\star(k)}$ (case 2), or reducing both relaxations by some small amount (case 3). Therefore, VCM-optimal relaxations are only applied to a single bound of a model-data constraint. It follows that the VCM can be rewritten in a more suggestive form,

$$\begin{aligned} \min_{x,\Delta,\delta} \quad & \|\Delta\|_1 + \|\delta\|_1 \\ \text{s.t.} \quad & L_i \leq M_i(x) - \Delta_i \leq U_i \qquad \text{for } i = 1, ..., N \\ & l_j \leq x_j - \delta_j \leq u_j \qquad \text{for } j = 1, ..., n. \end{aligned} \tag{7.5}$$

where sign of the relaxation determines to which bound it is applied. Therefore, the VCM can also be interpreted as providing the minimal additive "correction" to the model and parameters required to achieve consistency. This form shares structural similarities with the inclusion of an additive bias or discrepancy term in the probabilistic approach of [10].

An additional consequence of the VCM formulation is that the relaxed constraints will always be met with equality. This can again be demonstrated by contradiction. Continuing with the notation above, suppose that the constraints with upper bound relaxations are not always met with equality. Then there exists a $z \in \mathcal{F}'$ and an index k such that either $L_k \leq M_k(z) < U_k$ or $U_k \leq M_k(z) < U_k + \Delta_U^{\star(k)}$. So, either the relaxation is unnecessary or it can be made smaller. A similar argument can be made for constraints with lower bound relaxations. We conclude that the relaxed constraints are met with equality for all parameters in $\mathcal{F}'$. In the context of a VCM analysis, $\mathcal{F}'$ is the feasible set that arises from directly implementing all of the determined relaxations. Thus, the resulting model predictions of the relaxed QOIs (Section 2.5) are exact (without uncertainty) and equivalent to the adjusted bound. In practice, $\mathcal{F}'$ may consist of a single point, suggesting no parametric uncertainty and hence no uncertainty in model predictions of other QOIs. Such a result must be regarded with caution. This newfound certainty comes from a purely mathematical source (i.e., the search for *minimal* relaxations) rather than from any physical considerations. For these reasons, we often interpret the optimal relaxations as simply flagging potentially problematic constraints.

In general, inequality relaxation is a component of constrained optimization and has been used in different settings, e.g., the sum of infeasibilities method [19, p. 580] and the elastic filter [24, p. 101]. The application to UQ was suggested as early as [35, p. 41]. The optimal value of the above problem is still difficult to compute as the nonconvexity of the

models has yet to be addressed. For quadratic surrogate models, the VCM can be cast as an NQCQP and the convex relaxation techniques described in Section 2.7 can be applied. The computational aspect as well as comparison with the SCM is discussed in detail in Section 7.7 .

7.3 Linear examples

In order to develop some perspective on this new consistency measure, we first investigated the technique with linear models. In this setting, the VCM can be cast as a linear program and the bracketing strategy described in Section 2.7 is unnecessary. The examples in this section demonstrate that although the above formulation is intuitive, it fails to guarantee certain key properties that one might wish to have. The findings herein motivate the inclusion of weights into the vector consistency framework for an additional layer of flexibility.

Consider the following specialization of the Equation (7.3) to linear models,

$$\begin{aligned} &\min_{x,\delta} \ \|\delta\|_1 \\ &\text{s.t. } \mathbf{A}x \leq b + \delta \end{aligned} \tag{7.6}$$

where $\mathbf{A} \in \mathbb{R}^{N\times n}$ and $b \in \mathbb{R}^N$. As in the original setup, if the constraint $\mathbf{A}x \leq b$ is feasible, the minimizer $\delta^\star$ will be the zero vector. Suppose we start with a feasible dataset $\mathbf{A}x \leq b$, corresponding to some underlying ground truth, and tighten the constraint to inconsistency using a scaling factor $\alpha > 0$ and error vector $e \in \{0,1\}^N$ such that $\{x : \mathbf{A}x \leq b_\alpha\} = \varnothing$ where $b_\alpha = b - \alpha e$. Note that as discussed in the preceding section, the introduced error can be either interpreted as an error in the bounds or as a bias in the model.

We pose the following question: as α becomes arbitrarily large, can we guarantee that Equation (7.6) eventually recovers the error, i.e., $\delta^\star = \alpha e$? Can we expect to recover the underlying "true" feasible set? Although this would be a desirable property, the answer turns out to be no. This is illustrated in the following counterexample.

Let $\mathbf{A} = \begin{bmatrix} 1.5 \\ -1 \end{bmatrix}$, $b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, and $e = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Hence, the error is introduced only to the first QOI bound. With these settings, Equation (7.6) can be rewritten as,

$$\begin{aligned} &\min_{x,\delta} \ \|\delta\|_1 \\ &\text{s.t. } \begin{bmatrix} 1.5 \\ -1 \end{bmatrix} x - \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix} \leq \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix}. \end{aligned} \tag{7.7}$$

The dataset can be shown inconsistent for any $\alpha > 2.5$. The optimal relaxation $\delta^\star$ can be characterized by the intersection of two regions,

$$\begin{aligned} K_\alpha &= \{y \in \mathbb{R}^2 : \ \exists \delta \text{ such that } \|\delta\|_1 \leq c^\star, \ \ y \leq \delta\} \\ G_\alpha &= \{y \in \mathbb{R}^2 : y = \mathbf{A}x - b + \alpha t, x \in \mathbb{R}\}, \end{aligned} \tag{7.8}$$

where $c^\star$ is the smallest value such that the intersection is nonempty (i.e., the solution of Equation (7.7)). This intersection is displayed in Figure 7.1 for two values of α. Although the error has been introduced along the y_1 axis, G_α will always meet K_α along the y_2 axis due to its slope. Thus, the VCM will always produce a relaxation to the second constraint despite the error being introduced in only the first constraint. Simply put, $\delta^\star \neq \alpha t$ for any $\alpha > 0$. Regardless of the magnitude of the error, the VCM will always select the wrong constraint.

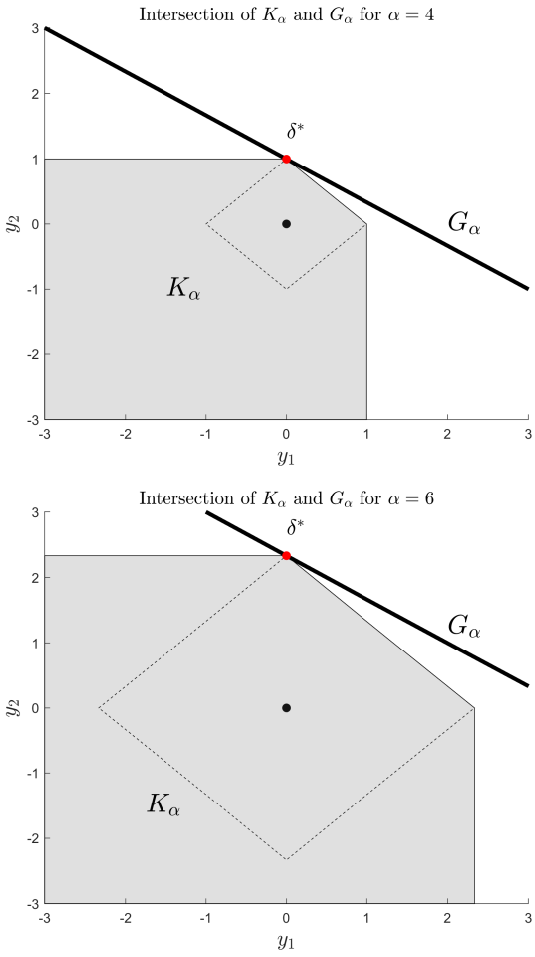

Figure 7.1: Illustration of the counter example. The dashed lines indicate the 1-norm ball with radius $c^\star$.

This single counterexample illustrates a somewhat contrived situation in which the VCM is unable to recover a specified error. To explore and test the significance of this result, we investigated the performance of the VCM on multiple random instances of Equation (7.6). How frequently is the technique unable to recover the ground truth?

Let $\mathbf{A} \in \mathbb{R}^{60\times 20}$ have entries selected from a uniform distribution on $[-1, 1]$, and let $b \in \mathbb{R}^{60}$ be chosen such that the inequality $\mathbf{A}x \leq b$ is feasible for some $x \in \mathbb{R}^{20}$. Additionally, let $e \in \{0, 1\}^{60}$ be defined as above with a total of n_e 1s placed in random entries. Notationally, n_e refers to the number of errors introduced. Let $\alpha > 0$ be such that $\{x : \mathbf{A}x \leq b_\alpha\}$ is empty. In order to measure the success of recovering the correct error, define the following ratios,

$$\begin{aligned} \phi_e &= \frac{|\text{supp}(\delta^\star) \cap \text{supp}(e)|}{n_e} \\ \phi_\delta &= \frac{|\text{supp}(\delta^\star) \cap \text{supp}(e)|}{|\text{supp}(\delta^\star)|} \end{aligned} \tag{7.9}$$

where $\text{supp}(a) = \{i : a_i \neq 0\}$ denotes the support of vector a, $|S|$ returns the cardinality of a set S, and $\delta^\star$ is the optimal value of Equation (7.6). The quantity ϕ_e measures the fraction of errors correctly identified by the optimization, whereas the term ϕ_δ measures the number of correctly identified errors relative to the total number of identifications. Both ratios range between zero and one, with $(\phi_e, \phi_\delta) = (1, 1)$ indicating perfect identification. If $\phi_e < 1$, then the approach identifies only some, but not all, of the errored constraints. If $\phi_\delta < 1$, then the algorithm identifies extra constraints. Together, these ratios account for both under-identifying and over-identifying the QOIs involved in the inconsistency. The results of $10,000$ random trials for $n_E = 1$ and two values of α are displayed in Figure 7.2.

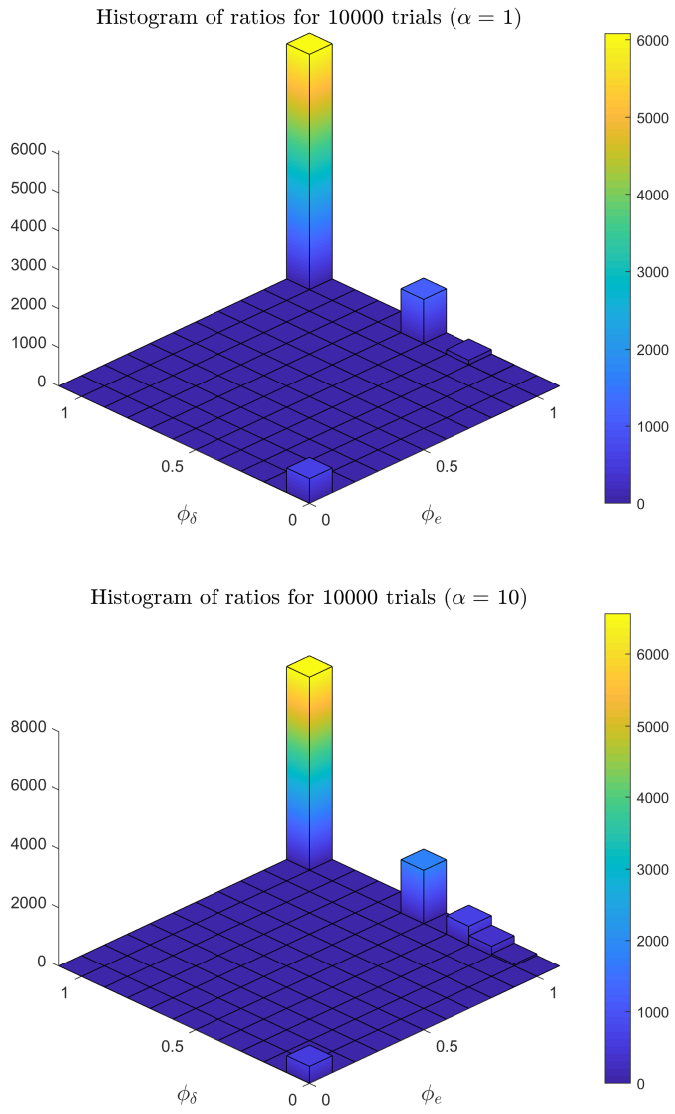

Figure 7.2: Histogram of ratios for 10,000 random trials with $n_e = 1$.

With this setup, the methodology displays exact identification for the vast majority of trials. When n_e is increased to $n_e = 4$, however, the number of possible explanations for the inconsistency increases as well, making the identification more challenging. This is illustrated below in Figure 7.3.

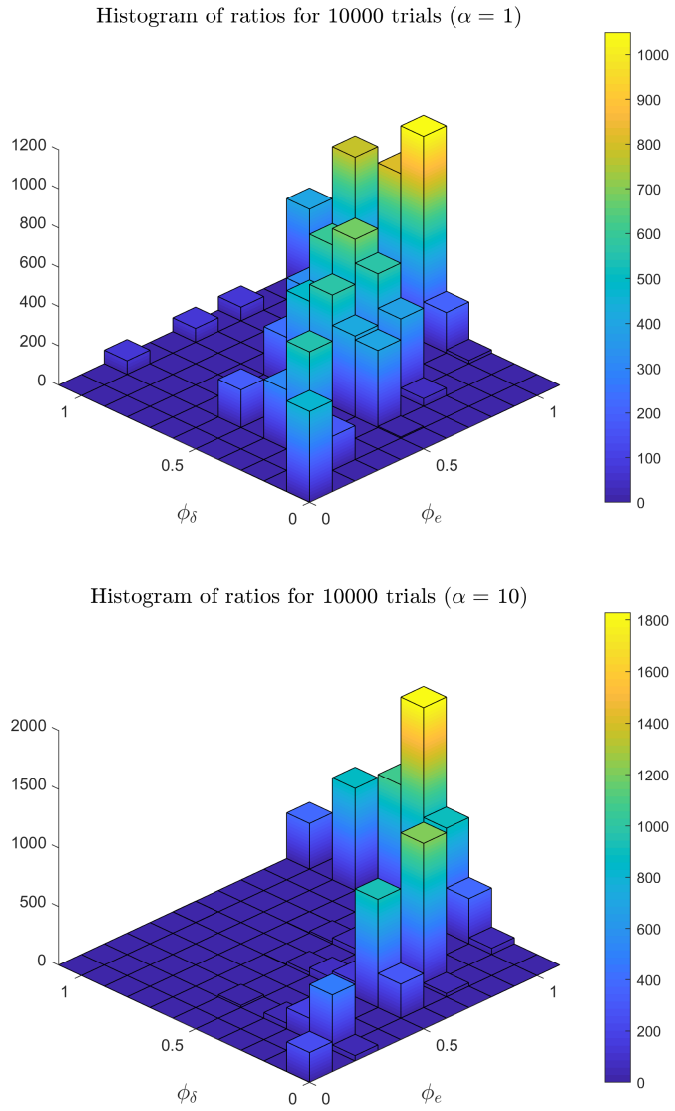

Figure 7.3: Histogram of ratios for 10,000 random trials with $n_e = 4$.

For a large magnitude error ($\alpha = 10$), the majority of trials yield $\phi_\delta \approx 0.5$ and $\phi_e = 1$. Thus, the VCM suggests relaxations to all four of the ground truth errored constraints plus approximately four extra constraints. In a lesser but significant number of trials, the VCM

is unable to identify all of the errors and suggests relaxations to extra constraints. There are many paths to consistency and in these cases Equation (7.6) finds alternative relaxation strategies (other than to the errors in e) that have smaller 1-norms.

In the context of resolving inconsistency, framing the problem as one of recovering a single "correct" relaxation is somewhat misrepresentative and oversimplifies the issue. As discussed above, for any given inconsistent dataset, there may be a multitude of relaxations that could lead to consistency. Each of these plausible relaxations may lead, after implementation, to different feasible parameter configurations. The presumption of sparsity, the guiding principle behind the VCM, does not alone provide enough information to pick out this so-called correct relaxation in a reliable manner. To achieve this would require physical insights and information from domain experts. This feature can be incorporated into the vector consistency framework through the inclusion of user-specified coefficients to the relaxations that act as weights.

The impact of weights can be illustrated by revisiting the earlier counterexample and solving a modified version of Equation (7.6):

$$\begin{aligned} &\min_{x,\delta} \ \|\delta\|_1 \\ &\text{s.t. } \mathbf{A}x \leq b + \mathbf{W}\delta, \end{aligned} \tag{7.10}$$

where $\mathbf{W}$ is a diagonal weighting matrix with nonnegative entries along the diagonal. A choice of $\mathbf{W}_{11} = 2$ and $\mathbf{W}_{22} = 1$ leads to the result in Figure 7.4, in which the errored constraint is now identified by the algorithm. The weights distort the 1-norm ball by stretching

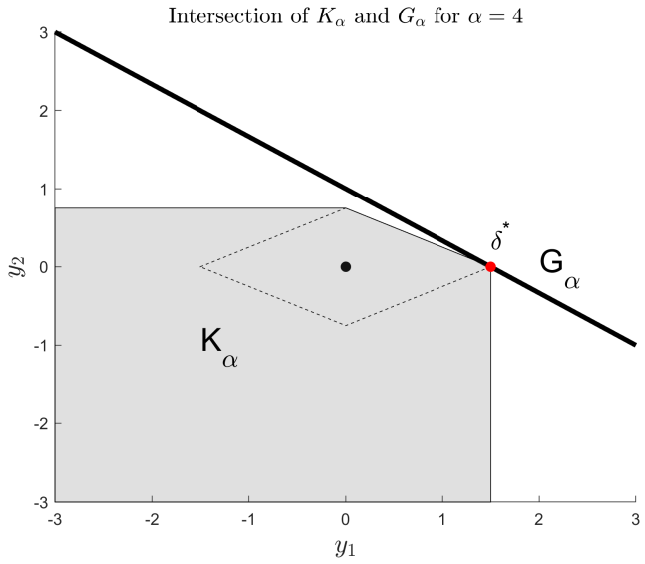

Figure 7.4: Counter example with weighted linear VCM.

along directions with large values. Hence, specifying a larger weight for relaxations to the

first constraint allows the error to be recovered. The real question becomes: "how to pick the appropriate weights?"

7.4 Weighted vector consistency

The inclusion of weights into the VCM, as described in the previous section for linear examples, adds an additional layer of flexibility to the framework. Specifically, it allows domain experts to emphasize or prioritize certain model-data constraints over others in the process of resolving an inconsistency. An example of this may be a situation in which one has prior knowledge or an opinion that certain experiments are more reliable than others. This notion of reliability may also be interpreted in the context of the underlying computational model; certain physics-based components of the model, perhaps linked to particular QOIs and therefore specific constraints, may be of higher fidelity than others. Naturally, how inconsistency is resolved should account for these preferences – we should be less inclined to adjust model-data constraints that are believed credible. For different problems, criteria other than reliability may be useful. For example, certain QOIs may be more relevant to the intended use of the dataset. In the context of predicion, QOIs that are most relevant to the prediction case should be prioritized in some fashion during the investigation. We extend the weighted linear VCM of Equation (7.10) to a more general form in Equation (7.11),

$$
\begin{aligned}
V_{\|\cdot\|_1}(w) = \min_{x,\Delta_L,\Delta_U,\delta_l,\delta_u} & \|\Delta_L\|_1 + \|\Delta_U\|_1 + \|\delta_l\|_1 + \|\delta_u\|_1 \\
\text{s.t.} \quad & L_i - W_L^{(i)}\Delta_L^{(i)} \le M_i(x) \le U_i + W_U^{(i)}\Delta_U^{(i)} \qquad \text{for } i = 1, ..., N \\
& l_j - w_l^{(j)}\delta_l^{(j)} \le x_j \le u_j + w_u^{(j)}\delta_u^{(j)} \qquad \text{for } j = 1, ..., n
\end{aligned}
\tag{7.11}
$$

where w collects the weight coefficients $\{W_L, W_U, w_l, w_u\}$. In the above formulation, the optimization prefers relaxing constraints with large weights as that action has a lesser impact on the penalty in the objective. In contrast, constraints with small weight coefficients are protected from relaxation as they require larger penalties in the objective. For instance, setting the parameter coefficients $w_l^{(j)}$ and $w_u^{(j)}$ to zero enforces the original parameter bounds and allows no modification. Hence, the corresponding relaxations $\delta_l^{(j)}$ and $\delta_u^{(j)}$ are forced to take on values of zero. Similarly, setting the weight coefficients of the ith model-data constraint, $W_L^{(i)}$ and $W_U^{(i)}$, to zero would imply that the corresponding constraint bounds are immutable. In essence, these weights specify the degree of flexibility afforded to certain constraint relaxations.

In general, the weights should be chosen to characterize the relative importance of certain model-data constraints over others. How to choose numerical values to quantify this relative importance, however, remains a challenge. There are several standard weight configurations that may be informative. For instance, the original vector consistency measure presented in Equation (7.3) is reconstructed by setting all weights to one. A list of simple weight settings and their interpretations are summarized in Table 7.1 for a generic constraint.

Table 7.1: Example weights for a generic constraint: $L - w_L\delta_L \leq f(x) \leq U + w_U\delta_U$

Weight	δ_L, δ_U — Interpretation
$w_L = w_U = 1$	Absolute change in bound ("*unit* weight")
$w_L = w_U = (U - L)$	$\delta_L + \delta_U \sim$ percent expansion in uncertainty interval ("*interval* weight")
$w_L = \lvert L\rvert, w_U = \lvert U\rvert$	Percent decrease/increase in lower/upper bound ("*bound* weight")
$w_L = w_U = 0$	No relaxation permitted ("*null* weight")

7.5 Using the VCM

The outcome of a VCM analysis can motivate several different actions. First, directly implementing the determined relaxations by expanding the bounds produces a consistent dataset. Second, the removal of the flagged QOIs, i.e., those associated with nonzero relaxations, leads to a consistent dataset composed of only the retained QOIs. Note, the removal of model-data constraints can be equivalently interpreted as widening the flagged QOI bounds beyond the recommended amount until the inequalities become inactive. Third, some combination of the first two strategies can be administered, leading to a consistent dataset with both expanded bounds and removed QOIs. In some cases, the inconsistency might be suspected to originate in the underlying computational model rather the experimental bounds. Then refining the underlying model, perhaps in a manner informed by the flagged QOIs and corresponding relaxations, may also lead to consistency. As discussed previously, the decision of how to proceed is to be resolved by domain scientists. Examples of these strategies in the context of a general consistency analysis are highlighted in the paragraph below.

In work by Feeley et al. [37], inadequacy in the computational model was primarily believed to be due to the uncertain parameter values rather than missing physics. The inconsistency of the dataset was rectified by identifying possibly problematic QOIs using the SCM and then revising the corresponding experimental bounds based on recommendations from the associated researchers. In Slavinskaya et al. [105], the temporary removal of certain QOIs was justified due to perceived faults in the instrument models rather than the underlying physical model. In Iavarone et al. [59], an inconsistent model was replaced by an updated model form consistent with experimental data.

Weights are included in the VCM framework to aid in this decision-making process. The VCM analysis can be repeated in a trial-and-error fashion, using different weight configurations to explore a given dataset's inconsistency. If the VCM detects the presence of a massive inconsistency, e.g., by always requiring the adjustment of numerous model-data constraints, then this suggests possibly severe limitations in the underlying model. The VCM and its weighted counterpart aim to provide a more efficient tool for examining inconsistency as

compared to the SCM. This is demonstrated in the following section where we revisit the GRI-Mech 3.0 and DLR-SynG examples.

7.6 Examples: GRI-Mech 3.0 and DLR-SynG

As with the SCM analysis described in Section 6.6 and Section 6.7, the GRI-Mech 3.0 and DLR-SynG datasets are used here as platforms to illustrate VCM-based resolutions to inconsistency. Note that both the results presented below and the previous SCM sensitivity analysis utilize the same type of redundant inequalities to strengthen the convex relaxation. Hence, the outcomes of both types of analysis can be directly compared. Of course, a key difference between both approaches is that the decision variable for the VCM is of greater dimension, i.e., $z = \begin{bmatrix} x^\intercal & \Delta_L^\intercal & \Delta_U^\intercal & \delta_l^\intercal & \delta_u^\intercal \end{bmatrix}^\intercal$. As the VCM is cast as a minimization, the computational strategy leads to VCM $\in$ [SDP, local solution]. Hence, a positive SDP result certifies that the dataset is inconsistent. Moreover, no collection of relaxations with 1-norm less than the SDP result can lead to consistency.

GRI-Mech 3.0

The GRI-Mech 3.0 dataset was previously demonstrated inconsistent by the SCM. The same is possible with the VCM; an initial application with *unit* weights on the QOIs and *null* weights on the parameters leads to $\mathrm{V}_{\|\cdot\|_1}(w) \in [0.017, 0.024]$. The positive lower bound of 0.017 proves that the dataset is inconsistent. Moreover, any collection of relaxations with 1-norm less than 0.017 will not result in consistency. Conversely, the upper bound of 0.024 corresponds to a local solution to the optimization. In this case, the following QOIs are identified by the analysis: the QOI #37 lower bound ("f5"), relaxed by 0.013 (a 0.8% decrease), and the QOI #36 upper bound ("f4"), relaxed by 0.011 (a 0.8% increase as well). As expected, these flagged QOIs are identical to those in the SCM analysis of Section 6.6.

The same analysis can be repeated for both *bound* and *interval* weights on the QOIs. For bound weights, $\mathrm{V}_{\|\cdot\|_1}(w) \in [0.011, 0.016]$. As presented in Table 7.1, the choice of bound weights leads to Δ (and δ) being interpreted as the percent change in the corresponding bound. Hence, the SDP result of 0.011 informs us that the total sum of these percent changes cannot be smaller than 1.1%. In this case, only the QOI #37 lower bound is identified, with a relaxation of 0.025 (a 1.6% increase). We have again confirmed the finding that the GRI-Mech dataset can be made consistent by removing just a single QOI. Moreover, we now know by exactly how much the corresponding bound must be changed to achieve consistency. Alternatively, interval weights produces $\mathrm{V}_{\|\cdot\|_1}(w) \in [0.25, 0.41]$ and demonstrates that the GRI-Mech 3.0 dataset can also be made consistent by adjusting the upper bounds of both QOIs #36 and #37.

The relaxations in terms of percentage change in the corresponding bound are displayed below in Figure 7.5 for the different weight configurations.

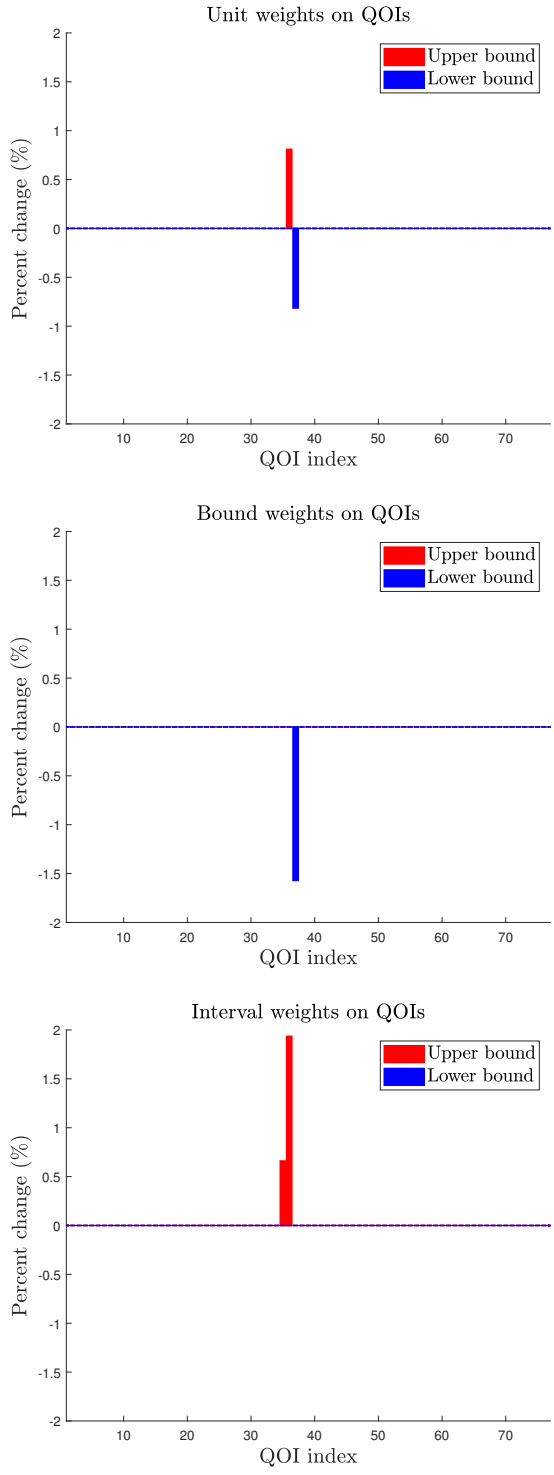

Figure 7.5: VCM optimal relaxations for three standard weight configurations.

These results illustrate that although the GRI-Mech 3.0 dataset is inconsistent, it is very close to consistency – expanding a single bound by under 2% provides a sufficient fix.

In each of the above results, QOI #37 was flagged either for the lower bound (with unit and bound weights) or the upper bound (with interval weights). The weights may be manipulated to ask more sophisticated questions about the inconsistency. For example, suppose QOI #37 and the associated uncertainties were believed accurately characterized. Then a natural question to ask is: "can the dataset be made consistent *without* adjusting QOI #37?" Setting $W_L^{(37)} = 0$ and fixing unit weights to the other QOI bounds leads to $\mathrm{V}_{\|\cdot\|_1}(w) \in [0.019, 0.35]$, with expansions to the upper bounds of QOIs #35 and #36 of 0.5% and 2% respectively. The answer to the question posed is therefore "yes, the dataset can be made consistent while preserving the contribution from QOI #37."

The initial result with unit weights flagged two bounds: the lower bound of QOI #37 and the upper bound of #36. We may further explore this particular set of relaxations by (a) fixing zero weights to the other bounds and (b) varying the weights for just the two identified bounds. This investigation produces the trade-off curve shown below in Figure 7.6.

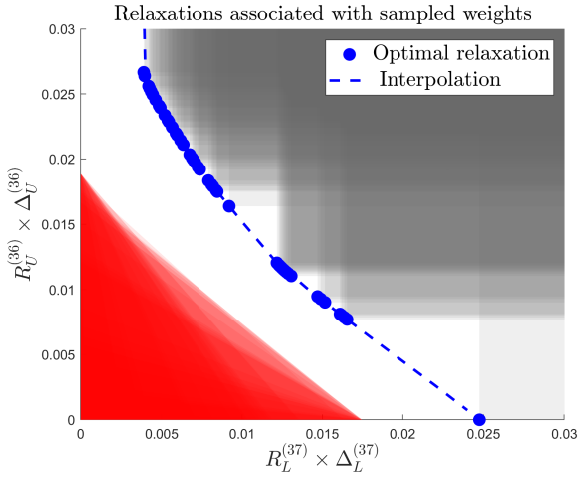

Figure 7.6: VCM relaxations for QOIs #36 and #36 (blue dots) due to random selection of weights. Other feasible relaxations lie in the shaded gray region. The red region is certified infeasible by the SDP results.

Another question that can be asked is, "can the dataset be made consistent by adjusting just the prior bounds on the parameters?". We note that this question is somewhat ill-considered as the surrogate models $M_i(x)$ are only constructed over $\mathcal{H}$ (or more accurately, in just the associated active parameters). In any case, attaching null weights to the QOIs and unit weights to the parameters produces $\mathrm{V}_{\|\cdot\|_1}(w) \in [0, 0.74]$. The lower bound of a single variable, x_{14}, is relaxed by 74%.

DLR-SynG

The SCM results of Section 6.7 suggested that many of the model-data constraints in the DLR-SynG dataset contribute to the inconsistency. In that example, the sensitivity-based iteration led to a roundabout procedure with several points of ambiguity. A VCM-based analysis, however, provides a more direct set of results.

An initial application of the VCM with unit weights on the QOI bounds and null weights on the parameter bounds produces $\mathrm{V}_{\|\cdot\|_1}(w) \in [7.15, 12.59]$. Hence, no collection of relaxations with 1-norm smaller than 7.15 can lead to consistency. In total, 43 QOI bounds are flagged with nonzero relaxations. These results are plotted below in Figure 7.7.

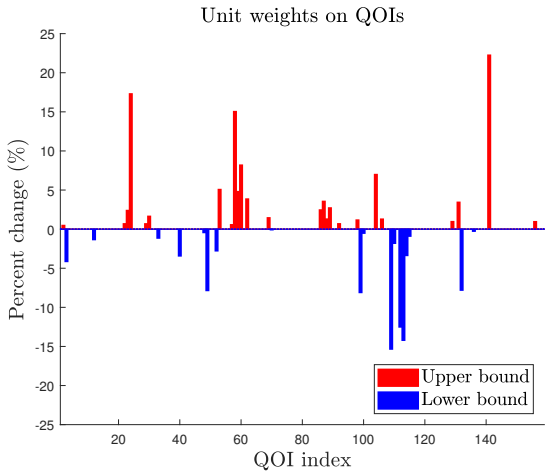

Figure 7.7: VCM relaxations for the DLR-SynG dataset with unit weights on the QOI bounds and null weights on the parameter bounds.

A number of the QOI bounds require expansion in excess of 10%. The scale of the inconsistency is clearly much greater than that of the GRI-Mech 3.0 dataset. As investigated in the previous example, alternative weight schemes can be used to further explore the inconsistency.

Placing bound weights on the QOIs and null weights on the parameters yields $\mathrm{V}_{\|\cdot\|_1}(w) \in [1.13, 1.69]$. Therefore, the total sum of the percent expansions in bounds cannot be smaller than 113%. As shown in Figure 7.8, this weight configuration leads to 37 constraint relaxations.

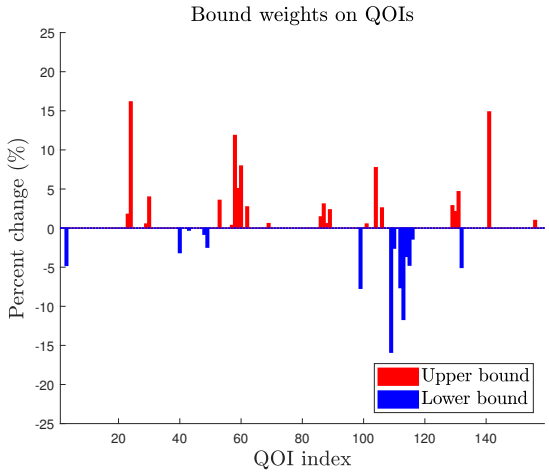

Figure 7.8: VCM relaxations for the DLR-SynG dataset with bound weights on the QOIs and null weights on the parameters.

Switching to interval weights for the QOIs leads to $\mathrm{V}_{\|\cdot\|_1}(w) \in [9.29, 13.22]$. In this case, 38 QOIs are identified as requiring bound relaxation.

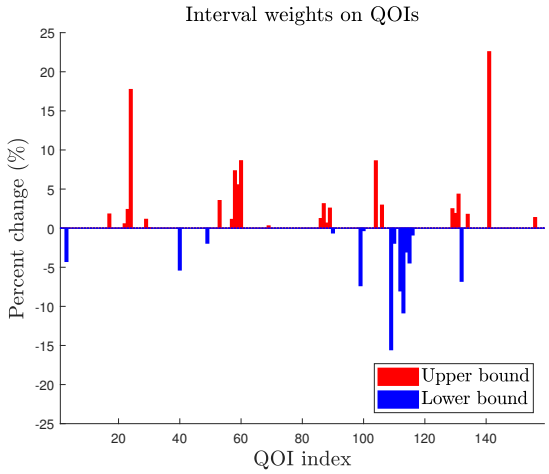

Figure 7.9: VCM relaxations for the DLR-SynG dataset with interval weights on the QOIs and null weights on the parameters.

As one might expect, there are many QOI bounds that commonly arise in the above results. For example, the upper bound of QOI #141 appears with significant relaxation regardless of the selected weight configuration. Suppose domain science determined that QOI #141 had both a highly reliable (surrogate) model and reliable experimental data. Then a pertinent question to ask is "can consistency be reached without having to relax either

bound of QOI #141?" Assigning null weights to both both bounds of the QOI and unit weights to the remaining QOIs produces $\mathrm{V}_{\|\cdot\|_1}(w) \in [8.05, 15.63]$, with 46 bound relaxations.

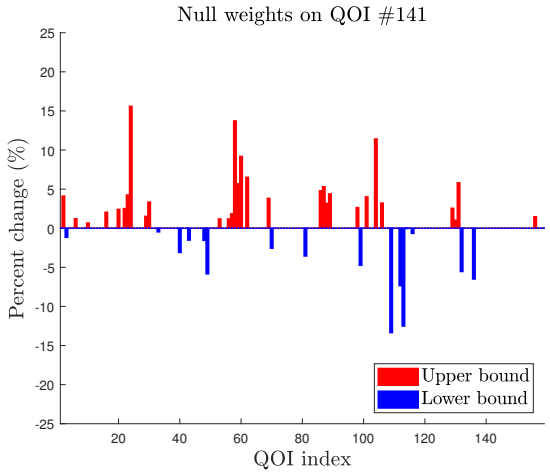

Figure 7.10: VCM relaxations for the DLR-SynG dataset with null weights on both QOI #141 as well as the parameters. The remaining QOIs were assigned unit weights.

Hence, the dataset can be made consistent without adjusting QOI #141.

7.7 Comparison with the SCM

The previous examples demonstrate that the VCM provides a more comprehensive tool for analyzing dataset inconsistency. Here, both the SCM and the VCM are treated in a unified setting to highlight their comparison. Note, the boldface text **SCM** and **VCM** are used below to refer to the specific formulations developed in this section.

For quadratic models, Equation (2.4) and Equation (7.11) can be rewritten in a more general form by rearranging terms and indices. A weighted version of the SCM can be reformulated as,

$$\begin{aligned} \textbf{SCM}: \quad \max_{x\in\mathbb{R}^n,\gamma} \quad & \gamma \\ \text{s.t.} \quad & \begin{bmatrix}1\\x\end{bmatrix}^\intercal \mathbf{Q}_i \begin{bmatrix}1\\x\end{bmatrix} \leq -w_i\gamma \qquad i = 1,...,N \\ & a_j^\intercal \begin{bmatrix}1\\x\end{bmatrix} \leq 0 \qquad j = 1,...,m \end{aligned} \tag{7.12}$$

where $\mathbf{Q}_i$ is the coefficient matrix of the corresponding surrogate model and the experimental bounds have been absorbed into the constant term, the parameter prior $\mathcal{H}$ is defined by a

collection of affine constraints constraints involving vectors a_j, and the weights $w_i \geq 0$. The VCM can be similarly rewritten as,

$$\begin{aligned} \textbf{VCM}: \quad \min_{x\in\mathbb{R}^n, \Delta\in\mathbb{R}^N, \delta\in\mathbb{R}^m} \quad & \mathbf{1}_N^\intercal \Delta + \mathbf{1}_m^\intercal \delta \\ \text{s.t.} \quad & \begin{bmatrix}1\\x\end{bmatrix}^\intercal \mathbf{Q}_i \begin{bmatrix}1\\x\end{bmatrix} \leq W_i \Delta_i \qquad i = 1, ..., N \\ & a_j^\intercal \begin{bmatrix}1\\x\end{bmatrix} \leq w_j \delta_j \qquad j = 1, ..., m \\ & \Delta \geq 0 \\ & \delta \geq 0 \end{aligned} \tag{7.13}$$

where $\mathbf{1}_d$ is the $d-$dimensional vector containing only ones as elements.

The SDP relaxations to both the **SCM** and **VCM** can be formulated in an equivalent manner. For the **SCM**, first recall that the constraints in Equation (7.12) can be rewritten as,

$$\begin{aligned} &\text{trace}\left(\begin{bmatrix}\mathbf{Q}_i & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}1\\x\\\gamma\end{bmatrix}\begin{bmatrix}1\\x\\\gamma\end{bmatrix}^\intercal\right) + \text{trace}\left(\begin{bmatrix}0 & 0 & 0.5w_i\\ 0 & 0 & 0\\ 0.5w_i & 0 & 0\end{bmatrix}\begin{bmatrix}1\\x\\\gamma\end{bmatrix}\begin{bmatrix}1\\x\\\gamma\end{bmatrix}^\intercal\right) \leq 0 \\ &\begin{bmatrix}\mathbf{A}^\intercal & 0\end{bmatrix}\begin{bmatrix}1\\x\\\gamma\end{bmatrix} \leq 0 \end{aligned} \tag{7.14}$$

where $\mathbf{A}$ is the matrix with columns $a_j \in \mathbb{R}^{n+1}$, i.e., $\mathbf{A} = \begin{bmatrix}a_1 & a_2 & ... & a_m\end{bmatrix} \in \mathbb{R}^{(n+1)\times m}$. In what follows, we use entrywise notation similar to `MATLAB` to clarify the subvectors and submatrices: $v(k{:}l)$ denotes the subvector containing the kth through lth entries of v and $\mathbf{V}(k{:}l, p{:}q)$ denotes the corresponding rectangular submatrix of $\mathbf{V}$. As in Section 2.7, $\mathcal{S}^d \subset \mathbb{R}^{d\times d}$ is used to denote the collection of symmetric matrices. Let $v = \begin{bmatrix}x^\intercal & \gamma\end{bmatrix}^\intercal$ denote the decision variable associated with **SCM**. The SDP relaxation can then be expressed as

$$\begin{aligned} \textbf{rSCM:} \quad \max_{\mathbf{Z}\in\mathcal{S}^{2+n}} \quad & \mathbf{Z}(2{+}n, 1) \\ \text{s.t.} \quad & \text{trace}\left(\mathbf{Q}_i \mathbf{Z}(1{:}1{+}n, 1{:}1{+}n)\right) + w_i \mathbf{Z}(2+n, 1) \leq 0 \qquad i = 1, ..., N \\ & \mathbf{A}^\intercal \mathbf{Z}(1{:}1{+}n, 1) \leq 0 \\ & \mathbf{A}^\intercal \mathbf{Z}(1{:}1{+}n, 1{:}1{+}n)\mathbf{A} \geq 0 \\ & \mathbf{Z}(1,1) = 1 \\ & \mathbf{Z} \succeq 0 \end{aligned} \tag{7.15}$$

Note that in the above formulation, we allow all of the redundant inequalities discussed in Section 6.5 to be included.

An SDP relaxation of the **VCM** is constructed in an identical fashion, however, with the decision variable of **VCM** set as $v = \begin{bmatrix}x^\intercal & \Delta^\intercal & \delta^\intercal\end{bmatrix}^\intercal$. For notational convenience, let

$d = 1{+}n{+}N{+}m$ and denote by $\text{diag}(w)$ the diagonal matrix with the vector w along the diagonal. The constraints in Equation (7.13) can be rewritten as,

$$\text{trace}\left(\begin{bmatrix}\mathbf{Q}_i & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}1\\ x\\ \Delta\\ \delta\end{bmatrix}\begin{bmatrix}1\\ x\\ \Delta\\ \delta\end{bmatrix}^{\intercal}\right) + \text{trace}\left(\begin{bmatrix}0 & 0 & -0.5W_ie_i^{\intercal} & 0\\ 0 & 0 & 0 & 0\\ -0.5W_ie_i & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}1\\ x\\ \Delta\\ \delta\end{bmatrix}\begin{bmatrix}1\\ x\\ \Delta\\ \delta\end{bmatrix}^{\intercal}\right) \leq 0$$

$$\begin{bmatrix}\mathbf{A}^{\intercal} & 0 & -\text{diag}(w)\end{bmatrix}\begin{bmatrix}1\\ x\\ \Delta\\ \delta\end{bmatrix} \leq 0 \tag{7.16}$$

where e_i refers to the ith standard basis vector. The SDP relaxation of **VCM** is therefore,

$$\begin{aligned}
\textbf{rVCM:}\quad \min_{\mathbf{Y}\in\mathcal{S}^d}\quad & \mathbf{1}_{N+m}^{\intercal}\mathbf{Y}(2{+}n{:}d, 1)\\
\text{s.t.}\quad & \text{trace}\left(\mathbf{Q}_i\mathbf{Y}(1{:}1{+}n, 1{:}1{+}n)\right) - W_i\mathbf{Y}(1{+}n{+}i, 1) \leq 0 \qquad i = 1, ..., N\\
& \mathbf{A}^{\intercal}\mathbf{Y}(1{:}1{+}n, 1) - \text{diag}(w)\mathbf{Y}(2{+}n{+}N{:}d, 1) \leq 0\\
& \mathbf{A}^{\intercal}\mathbf{Y}(1{:}1{+}n, 1{:}1{+}n)\mathbf{A} - \text{diag}(w)\mathbf{Y}(2{+}n{+}N{:}d, 1{:}1{+}n)\mathbf{A} -\\
& \qquad \mathbf{A}^{\intercal}\mathbf{Y}(1{:}1{+}n, 2{+}n{+}N{:}d)\text{diag}(w) +\\
& \qquad \text{diag}(w)\mathbf{Y}(2{+}n{+}N{:}d, 2{+}n{+}N{:}d)\text{diag}(w) \geq 0\\
& \mathbf{Y}(2{+}n{:}d, 1) \geq 0\\
& \mathbf{Y}(1, 1) = 1\\
& \mathbf{Y} \succeq 0.
\end{aligned} \tag{7.17}$$

Due to the nature of the SDP relaxation, the following inequalities are immediate: **SCM** $\leq$ **rSCM** and **rVCM** $\leq$ **VCM**.

The GRI-Mech 3.0 and DLR-SynG examples of the previous section illustrated how different weighting configurations can be used to explore different resolutions to an inconsistency. For example, null weights were used to enforce certain constraints. In fact, null weights can also be helpful in proving inconsistency, which requires that **rVCM** > 0. To see this, consider two different VCM problems for the same dataset. Let $\mathbf{rVCM}_0$ denote the first VCM problem and let $\mathbf{rVCM}_1$ denote the second. Let both problems share identical weights except for those indices in the subsets $\mathcal{J}_W$ (QOI indices) and $\mathcal{J}_w$ (parameter indices). For these subsets, let $\mathbf{rVCM}_0$ have null weights, i.e., $i \in \mathcal{J}_W \Leftrightarrow W_i = 0$ and $j \in \mathcal{J}_w \Leftrightarrow w_j = 0$. Moreover, define $\mathcal{J} = \{1{+}n{+}\mathcal{J}_W\} \cup \{1{+}n{+}N{+}\mathcal{J}_w\}$ to be the adjusted set of indices to account for the ordering in the decision variable. Suppose $\mathbf{Y}_0^{\star}$ is the minimizer of $\mathbf{rVCM}_0$. Take $\mathbf{Y} \in \mathcal{S}^d$ to be

$$\mathbf{Y}(i,j) = \begin{cases} 0 & i \in \mathcal{J} \text{ or } j \in \mathcal{J}\\ \mathbf{Y}_0^{\star}(i,j) & \text{otherwise.}\end{cases} \tag{7.18}$$

Hence, we are simply zeroing out the rows and columns associated with the null weights. Thus, $\mathbf{Y}$ is feasible with respect to $\mathbf{rVCM}_0$ and hence $\mathbf{1}_{N+m}^\intercal \mathbf{Y}(2{+}n{:}d, 1) = \mathbf{1}_{N+m}^\intercal \mathbf{Y}_0^\star(2{+}n{:}d, 1)$. Moreover, note that $\mathbf{Y}$ is also feasible with respect to $\mathbf{rVCM}_1$ thus implying that $\mathbf{rVCM}_1 \leq \mathbf{rVCM}_0$. This was numerically illustrated in the previous section. For example, included a null weight on one of the QOIs in the GRI-Mech 3.0 dataset led to the SDP relaxation increasing from 0.017 to 0.019. In other cases, if $\mathbf{rVCM}_1$ cannot prove inconsistency, computing $\mathbf{rVCM}_0$ may do so.

Given the success of the VCM in handling inconsistent datasets, a natural question to ask is "can the VCM supplant the SCM?" We answer this question in the negative by demonstrating that **rSCM** is a stronger tool for detecting inconsistency.

Theorem 7.7.1. *If* $\boldsymbol{rSCM} \geq 0$*, then* $\boldsymbol{rVCM} = 0$ *for any choice of weights.*

Proof. Suppose $\mathbf{rSCM} \geq 0$ (the dataset is provably consistent) and let $\mathbf{Z}^\star \in \mathcal{S}^{2+n}$ be the corresponding maximizer. Let $W \geq 0$ and $w \geq 0$ be any weight configuration. Define $\mathbf{Y} \in \mathcal{S}^d$ by

$$\mathbf{Y}(i,j) = \begin{cases} \mathbf{Z}^\star(i,j) & i,j \leq 1{+}n \\ 0 & \text{otherwise} \end{cases} \tag{7.19}$$

Since $\mathbf{Z}^\star$ is feasible with respect to Equation (7.15) and $\mathbf{Z}^\star(2+n, 1) \geq 0$, $\mathbf{Y}$ must be feasible with respect to Equation (7.17). Moreover, $\mathbf{1}_{N+m}^\intercal \mathbf{Y}(2{+}n{:}d, 1) = 0$ and since **rVCM** is always nonnegative, $\mathbf{rVCM} = 0$. □

The contrapositive of Theorem 7.7.1 states that if $\mathbf{rVCM} > 0$ for some choice of weights, then $\mathbf{rSCM} < 0$. Hence, the collection of datasets for which **rVCM** proves inconsistency is a subset of the collection of datasets for which **rSCM** proves inconsistency.

It turns out that this containment is strict. We demonstrate this by constructing a numerical example in which $\mathbf{rSCM} < 0$ (SCM proves inconsistency) and $\mathbf{rVCM} = 0$ (VCM does not prove inconsistency). Consider a purely artificial dataset with dimensions $n = 2$, $N = 2$, and $m = 2$.

$$\begin{aligned} Q_1 &= \begin{bmatrix} 0.0881 & 0.460 & 0.4769 \\ 0.460 & 0.5613 & 0.4948 \\ 0.4769 & 0.4948 & 0.3550 \end{bmatrix} & Q_2 &= \begin{bmatrix} 0.2448 & 0.1876 & 0.1492 \\ 0.1876 & 0.2664 & 0.7218 \\ 0.1492 & 0.7218 & 0.1476 \end{bmatrix} \\ a_1 &= \begin{bmatrix} 0.9207 \\ 0.9295 \\ 0.1368 \end{bmatrix} & a_2 &= \begin{bmatrix} 0.8716 \\ 0.0124 \\ 0.7220 \end{bmatrix} \\ W_1 &= W_2 = w_1 = w_2 = 1 \end{aligned} \tag{7.20}$$

Here, the only redundant constraint added to the **rSCM** is $a_1^\intercal \mathbf{Z}(1{:}3, 1{:}3) a_2 \geq 0$ as the other constraints are implied by $\mathbf{Z} \succeq 0$, and similarly for **rVCM**. Using CVX [52, 53], it can be shown that $\mathbf{rSCM} = -1.0857$ and $\mathbf{rVCM} = 0$. To avoid issues of precision in the numerical evaluation of **rVCM**, we round the minimizer $Y^\star$ to the second digit and find that it is still

feasible with some margin. Hence, **rVCM** $= 0$ and **rSCM** < 0 therefore implying that the converse of Theorem 7.7.1 does not hold.

7.8 An extension to rational quadratic surrogates

A natural follow-up question to the previous sections is: can these developments be directly extended to datasets composed of rational quadratic (RQ) surrogate models? It turns out that the extension is not so immediate. To illustrate this, we first review how the SCM is computed for such datasets.

Consider the following version of the SCM, without the coefficient to γ,

$$\begin{aligned} C_{\mathrm{D}} := \max_{\gamma,\ x\in\mathcal{H}} \quad & \gamma \\ \text{s.t.} \quad & L_i + \gamma \leq \frac{N_i(x)}{D_i(x)} \leq U_i - \gamma \\ & \text{for } i = 1, \ldots, N. \end{aligned} \tag{7.21}$$

where $N_i(x) = \begin{bmatrix}1\\ x\end{bmatrix}^{\mathsf{T}} \mathbf{N}_i \begin{bmatrix}1\\ x\end{bmatrix}$ and $D_i(x) = \begin{bmatrix}1\\ x\end{bmatrix}^{\mathsf{T}} \mathbf{D}_i \begin{bmatrix}1\\ x\end{bmatrix}$ are known RQ surrogate models fit using the strategies discussed in Chapter 4.

Recall that for such surrogates constructed over $\mathcal{H}$, we impose the condition that $1 \leq D_i(x) \leq \kappa_i$ for all $x \in \mathcal{H}$. Since the denominator is always positive in this region, Equation (7.21) may be rewritten as,

$$\begin{aligned} C_{\mathrm{D}} := \max_{\gamma,\ x\in\mathcal{H}} \quad & \gamma \\ \text{s.t.} \quad & \begin{bmatrix}1\\ x\end{bmatrix}^{\mathsf{T}} (\mathbf{N}_i + (\gamma - U_i)\mathbf{D}_i) \begin{bmatrix}1\\ x\end{bmatrix} \leq 0 \\ & \begin{bmatrix}1\\ x\end{bmatrix}^{\mathsf{T}} ((L_i + \gamma)\mathbf{D}_i - \mathbf{N}_i) \begin{bmatrix}1\\ x\end{bmatrix} \leq 0 \\ & \text{for } i = 1, \ldots, N. \end{aligned} \tag{7.22}$$

where we have expanded the quadratics to emphasize the cubic interaction between the scalar decision variable γ and the components quadratic in the vector decision variable x. Importantly, this means that the formulation is no longer an NQCQP. Thus the strategies discussed in Section 2.7 for computing an upper bound $\overline{C}_D$ through convex relaxation cannot be directly applied. Note that for fixed γ, the constraints in Equation (7.22) (including the prior constraint $x \in \mathcal{H}$) can be assembled into a quadratic "pseudo" dataset. To handle the difficulty, Feeley [35, Section 5.4] proposed a bisection-on-γ algorithm in which each step fixes a value for γ, formulates the corresponding quadratic pseudo-dataset, checks the feasibility of that dataset (i.e., computes the standard quadratic consistency measure), adjusts γ, and then iterates. This strategy cannot be directly extended to the convex relaxation of the

VCM. Recall in the VCM, the scalar variable γ is replaced by the vector variables Δ_L and Δ_U. Hence, in the resulting formulation there are multiple cubic interactions among the different $\Delta^{(i)}$ and components of x.

For RQ surrogate models, how the relaxations are implemented becomes important if one wishes to use B2BDC's computational toolbox. As discussed above, the standard way of incorporating the relaxations, i.e.,

$$L_i - \Delta_L^{(i)} \leq \frac{N_i(x)}{D_i(x)} \leq U_i + \Delta_U^{(i)}, \tag{7.23}$$

cannot be implemented using the bisection approach associated with the SCM as each bound has a unique relaxation variable $\Delta^{(i)}$. Note that we may still solve this problem to get a local solution, and hence an upper bound on the VCM. Our aim, however, is to find a useful lower bound too. An alternative approach is to formulate the relaxations by the pairing,

$$\begin{aligned} N_i(x) - U_i D_i(x) &\leq \Delta_U^{(i)} \\ L_i D_i(x) - N_i(x) &\leq \Delta_L^{(i)}. \end{aligned} \tag{7.24}$$

Note that for notational convenience we ignore weights as well as the relaxations to the parameter bounds. As the QOI relaxations (Δ_L, Δ_U) appear linearly with the vector decision variable x, there is no difficulty in applying the standard B2BDC SDP techniques. The two formulations in Equation (7.23) and Equation (7.24), however, have different meanings. For instance,

$$\left.\begin{cases} N_i(x) - U_i D_i(x) \leq \Delta_U^{(i)} \\ L_i D_i(x) - N_i(x) \leq \Delta_L^{(i)} \end{cases}\right\} \Leftrightarrow L_i - \frac{1}{D_i(x)}\Delta_L^{(i)} \leq \frac{N_i(x)}{D_i(x)} \leq U_i + \frac{1}{D_i(x)}\Delta_U^{(i)}. \tag{7.25}$$

The relaxation in this alternate formulation essentially has a parameter dependent weighting $(D_i^{-1}(x))$ associated with it. Since $1 \leq D_i(x) \leq \kappa_i$ for $x \in \mathcal{H}$, the following general relations hold:

$$L_i - \frac{1}{D_i(x)}\Delta_L^{(i)} \leq \frac{N_i(x)}{D_i(x)} \leq U_i + \frac{1}{D_i(x)}\Delta_U^{(i)} \Rightarrow L_i - \Delta_L^{(i)} \leq \frac{N_i(x)}{D_i(x)} \leq U_i + \Delta_U^{(i)}, \tag{7.26}$$

and conversely,

$$L_i - \frac{1}{\kappa_i}\Delta_L^{(i)} \leq \frac{N_i(x)}{D_i(x)} \leq U_i + \frac{1}{\kappa_i}\Delta_U^{(i)} \Rightarrow L_i - \frac{1}{D_i(x)}\Delta_L^{(i)} \leq \frac{N_i(x)}{D_i(x)} \leq U_i + \frac{1}{D_i(x)}\Delta_U^{(i)}. \tag{7.27}$$

The implication in Equation (7.26) says that the outcome of solving the associated VCM problem, i.e.,

$$\begin{aligned} \textbf{VCM-RQ}' := \min_{x,\Delta_L,\Delta_U} \; & \|\Delta_L\|_1 + \|\Delta_U\|_1 \\ \text{s.t.} \quad & N_i(x) - U_i D_i(x) \leq \Delta_U^{(i)} \quad \text{for } i = 1, ..., N \\ & L_i D_i(x) - N_i(x) \leq \Delta_L^{(i)} \quad \text{for } i = 1, ..., N \\ & l_j \leq x_j \leq u_j \quad \text{for } j = 1, ..., n \end{aligned} \tag{7.28}$$

leads to a feasible bound relaxation. If we let **VCM-RQ** denote the solution to the VCM with constraints of the form in Equation (7.23), then this implies **VCM-RQ**$'$ $\geq$ **VCM-RQ**. Although it differs from the typical VCM construction, the formulation **VCM-RQ**$'$ provides one strategy for extending the VCM to RQ surrogates. In general, the impact of parameter dependent weighting, as in Equation (7.25) would be a interesting area of future investigation.

7.9 Consistency measures that count: a possible future direction

The VCM strategies discussed and illustrated in the preceding sections are both practical and useful additions to the B2BDC framework. They are not, however, without flaws. The original motivation was to address the question posed by Equation (7.2): "what is the fewest number of constraint modifications required to reach consistency?" Switching to the 1-norm was a natural step in formulating a more tractable question: "what is the smallest sum of independent constraint relaxations required to reach consistency?" This new question became the baseline for the VCM. By definition, the SDP relaxation provides a global lower bound on this objective, with two interpretations: first, as a proof of inconsistency; and second, as a statement about the scale of the inconsistency. Investigating only the sum of relaxations, however, does not provide provable information on how the relaxations are distributed among the dataset constraints. For this, we relied on local solutions. In this section, we briefly illustrate how VCM-like ideas can be used to forge additional tools that provide complementary information. Although the results here are comparable to those in the preceding sections, we believe that further investigations in this direction could be fruitful in developing new consistency measures.

An alternative formulation of Equation (7.2) with the aim of explicitly counting constraint removals can be constructed by casting the relaxations as binary decision variables,

$$\begin{aligned} V_{\|\cdot\|_0} = \min_{x,\Delta} \; & \sum_{i=1}^{N} \Delta_i \\ \text{s.t. } & L_i - \infty \cdot \Delta_i \leq M_i(x) \leq U_i + \infty \cdot \Delta_i \\ & \Delta_i \in \{0,1\} \quad \text{for } i = 1, ..., N \\ & l_j \leq x_j \leq u_j \quad \text{for } j = 1, ..., n. \end{aligned} \tag{7.29}$$

Of course this cannot be implemented directly, but the intention is clear – a value of $\Delta_i = 1$ implies the corresponding model-data constraint is relaxed until ineffective, which is the same as removing the constraint. Hence, the sum in the objective counts the number of these relaxations/removals. The choice of ∞ is rather excessive and can be replaced by any finite quantity that renders the constraint inactive. One way of selecting such values is

through prediction over the prior $\mathcal{H}$. Let $B_L^{(i)}$ and $B_U^{(i)}$ be such that,

$$\begin{aligned} B_L^{(i)} &\geq L_i - \min_{x \in \mathcal{H}} M_i(x) \\ B_U^{(i)} &\geq \max_{x \in \mathcal{H}} M_i(x) - U_i. \end{aligned} \tag{7.30}$$

For quadratic or polynomial surrogate models, we may compute the values of B_L and B_U by replacing the predictions over $\mathcal{H}$ with their SDP relaxations. This ensures that the values are larger than necessary to render the constraints inactive over the prior. Equation (7.32) may then be rewritten as,

$$\begin{aligned} V_{\|\cdot\|_0} = \min_{x,\Delta} \; & \sum_{i=1}^{N} \Delta_i \\ \text{s.t. } & L_i - B_L^{(i)} \Delta_i \leq M_i(x) \leq U_i + B_U^{(i)} \Delta_i \\ & \Delta_i \in \{0,1\} \quad \text{for } i = 1, ..., N \\ & l_j \leq x_j \leq u_j \quad \text{for } j = 1, ..., n. \end{aligned} \tag{7.31}$$

Relaxing the binary constraints from $\Delta_i \in \{0,1\}$ to the interval constraint $\Delta_i \in [0,1]$ leads to a continuous optimization problem and ensures that optimal value is not greater than the optimal value of Equation (7.31). We term this new formulation as the *count* VCM, or $V_\#$.

$$\begin{aligned} V_\# := \min_{x,\Delta} \; & \sum_{i=1}^{N} \Delta_i \\ \text{s.t. } & L_i - \Delta_i B_L^{(i)} \leq M_i(x) \leq U_i + \Delta_i B_U^{(i)} \\ & \Delta_i \in [0,1] \quad \text{for } i = 1, ..., N \\ & l_j \leq x_j \leq u_j \quad \text{for } j = 1, ..., n. \end{aligned} \tag{7.32}$$

The SDP relaxation of Equation (7.32) therefore provides a lower bound on the above problem, which is in turn a lower bound on the number of constraint removals required to reach consistency. We additionally note that this new formulation is essentially a version of the weighted VCM in Equation (7.11), albeit with additional constraints and a special choice of weights.

Applying this new tool to the DLR-SynG dataset produces somewhat weak results. The computational strategy finds that $V_\# \in [4.15, 5.34]$, certifying that fewer than 5 QOI removals can never bring about consistency. The corresponding relaxations Δ are plotted below in Figure 7.11.

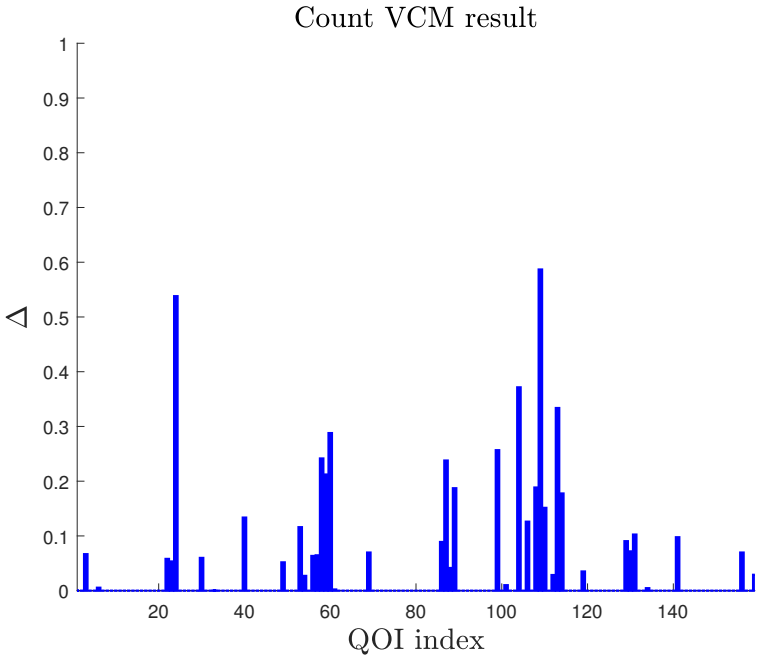

Figure 7.11: Count VCM relaxations.

Rounding up the nonzero results in Figure 7.11 to one and summing produces an upper bound on Equation (7.31). Hence the removal of the 40 identified QOIs (flagged with $\Delta_i > 0$) leads to a consistent dataset. Hence, $V_{\|\cdot\|_0} \in [5, 40]$. Although this result is not as informative as one might hope, it does provide additional information that the previous VCM strategies did not. Even though the computation of $V_{\#}$ is quite tight, the wide gap in $V_{\|\cdot\|_0}$ is due to relaxing the binary constraint.

7.10 Chapter summary

Consistency analysis is a fundamental component of B2BDC. In this chapter, a new consistency measure, termed the *vector consistency measure* (VCM), was introduced to address shortcomings of the SCM in diagnosing inconsistency. The utility of the VCM was demonstrated in two real-world cases arising from combustion chemistry: GRI-Mech 3.0 and DLR-SynG. In both of these examples, the VCM led to new insights and enabled new questions to be asked of an inconsistency. We note that the material in these sections is based on and was first reported in our recent work [54].

The end of the chapter saw both the comparison of existing methods as well as a suggestion for possible future directions. The SCM was shown to be stronger than the VCM in proving consistency whereas the VCM was demonstrated to be more practical for diagnosing inconsistency. New consistency measures that provide complementary information, such as

improvements to the count VCM, would be a welcome addition to the B2BDC framework. These final sections highlight the following important fact: there is no universal consistency measure. Rather, we see that a more complete picture of an inconsistent dataset requires a toolkit capable of exploring different questions. In this spirit, we emphasize that consistency measures are numerical tools to assist domain scientists.

Chapter 8

Conclusion

The methodological contributions of this dissertation are twofold. First, Chapter 3 presented an iterative approach for dataset construction that progressively hones in on useful regions of the parameter space. This new strategy was a direct outcome of exploring the relationship between B2BDC and Bayesian history matching. These techniques were extended in Chapter 4 and Chapter 5. Second, a novel tool for diagnosing inconsistent datasets was motivated in Chapter 6 and developed in Chapter 7. This tool, the vector consistency measure (VCM), investigates inconsistent datasets by seeking out the fewest number of constraint modifications that lead to consistency. The inclusion of weights allow domain scientists to play a more involved role in the resolution process.

These two contributions essentially tackle the same question: how can model-data agreement be quantified while accounting for both parametric and experimental uncertainty? In B2BDC, this question is posed by first constructing a dataset and then verifying its consistency. The iterative strategy discussed above extends the method's applicability by performing consistency checks sequentially over a shrinking feasible region. Through this procedure, accurate representation of the dataset QOIs becomes paramount. At each iteration, fitting is focused on the corresponding feasible set and the surrogate models are continually improved as accurately-represented QOIs are assimilated into the analysis. The procedure terminates if all QOIs are accounted for or if inconsistency is proven. If a dataset is consistent, the analysis can shift to other vital tasks such as prediction. Inconsistency, however, stymies further progress on this front and demonstrates disagreement among models, data, and prior information. The VCM seeks to quantify this disagreement. Incorporating weights into the VCM framework facilitates a more detailed diagnosis of the inconsistency by exploring multiple ways of reconciling the conflict. For example, the presence of a massive inconsistency could indicate significant model deficiencies. The information gleaned from such analysis could then be used to aid in revising the underlying computational model and/or observations. In essence, these steps of dataset construction, consistency checks, and further diagnosis serve as a general workflow to accomplish model validation in the presence of uncertainty.

Bibliography

[1] I. Andrianakis, N. McCreesh, I. Vernon, T. J. McKinley, J. E. Oakley, R. N. Nsubuga, M. Goldstein, and R. G. White, "Efficient history matching of a high dimensional individual-based hiv transmission model", *SIAM/ASA Journal on Uncertainty Quantification*, vol. 5, no. 1, pp. 694–719, 2017. DOI: 10.1137/16M1093008.

[2] I. Andrianakis, I. R. Vernon, N. McCreesh, T. J. McKinley, J. E. Oakley, R. N. Nsubuga, M. Goldstein, and R. G. White, "Bayesian history matching of complex infectious disease models using emulation: A tutorial and a case study on hiv in uganda", *PLoS Comput. Biol.*, vol. 11, 2015. DOI: 10.1371/journal.pcbi.1003968.

[3] I. Andrianakis, I. Vernon, N. McCreesh, T. J. McKinley, J. E. Oakley, R. N. Nsubuga, M. Goldstein, and R. G. White, "History matching of a complex epidemiological model of human immunodeficiency virus transmission by using variance emulation", *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 66, no. 4, pp. 717–740, 2017. DOI: 10.1111/rssc.12198.

[4] K. M. Anstreicher, "Semidefinite programming versus the refomulation-linearization technique for nonconvex quadratically constrained quadratic programming", *Journal of Global Optimization*, vol. 43, pp. 471–484, 2009. DOI: 10.1007/s10898-008-9372-0.

[5] K. Anstreicher and H. Wolkowicz, "On lagrangian relaxation of quadratic matrix constraints", *SIAM Journal on Matrix Analysis and Applications*, vol. 22, no. 1, pp. 41–55, 2000. DOI: 10.1137/S0895479898340299.

[6] M. Ashraphijuo, S. Fattahi, J. Lavaei, and A. Atamtürk, "A strong semidefinite programming relaxation of the unit commitment problem", in *2016 IEEE 55th Conference on Decision and Control (CDC 2016)*, 2016, pp. 694–701. DOI: 10.1109/CDC.2016.7798349.

[7] N. Baker, F. Alexander, T. Bremer, A. Hagberg, Y. Kevrekidis, H. Najm, M. Parashar, A. Patra, J. Sethian, S. Wild, *et al.*, "Workshop report on basic research needs for scientific machine learning: Core technologies for artificial intelligence", USDOE Office of Science (SC), Washington, DC (United States), Tech. Rep., 2019.

[8] T. Başar, *Uncertainty in Complex Networked Systems: In Honor of Roberto Tempo*. Springer, 2018, ISBN: 9783030046309. DOI: 10.1007/978-3-030-04630-9.

[9] M. J. Bayarri, J. O. Berger, J. A. Cafeo, G. Garcia-Donato, F. Liu, J. Palomo, R. J. Parthasarathy, R. Paulo, J. Sacks, and D. Walsh, "Computer model validation with functional output", *Annals of Statistics*, vol. 35, no. 5, pp. 1874–1906, 2007. DOI: 10.1214/009053607000000163.

[10] M. J. Bayarri, J. O. Berger, R. Paulo, J. Sacks, J. A. Cafeo, J. Cavendish, C. H. Lin, and J. Tu, "A framework for validation of computer models", *Technometrics*, vol. 49, no. 2, pp. 138–154, 2007. DOI: 10.1198/004017007000000092.

[11] R. E. Bellman, *Adaptive control processes: a guided tour*. Princeton university press, 1961, ISBN: 9780691652214.

[12] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust Optimization*, ser. Princeton Series in Applied Mathematics. Princeton University Press, Oct. 2009, ISBN: 9780691143682.

[13] D. Bertsimas and J. Sethuraman, "Moment problems and semidefinite programming", in *Handbook on semidefinite programming: Theory, Algorithms, and Applications*, H. Wolkowicz, R. Saigal, and L. Vandenberghe, Eds., Springer, 2000, ch. 16, pp. 469–509, ISBN: 978-1-4615-4381-7. DOI: 10.1007/978-1-4615-4381-7.

[14] F. Blanchini and S. Miani, *Set-theoretic methods in control*. Springer, 2008, ISBN: 9783319179339. DOI: 10.1007/978-3-319-17933-9.

[15] N. Bliznyuk, D. Ruppert, C. Shoemaker, R. Regis, S. Wild, and P. Mugunthan, "Bayesian calibration and uncertainty analysis for computationally expensive models using optimization and radial basis function approximation", *Journal of Computational and Graphical Statistics*, vol. 17, no. 2, pp. 270–294, 2008. DOI: 10.1198/106186008X320681.

[16] G. E. P. Box and W. G. Hunter, "The experimental study of physical mechanisms", *Technometrics*, vol. 7, no. 1, pp. 23–42, 1965. DOI: 10.2307/1266125.

[17] G. E. P. Box, W. G. Hunter, and J. S. Hunter, *Statistics for Experimenters*, 2nd ed., ser. Wiley Series in Probability and Statistics. Hoboken, New Jersey: John Wiley & Sons, Inc., 2005, ISBN: 9780471718130.

[18] G. E. Box and N. R. Draper, *Empirical model-building and response surfaces*. John Wiley & Sons, 1987, ISBN: 9780471810339.

[19] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004, ISBN: 9780511804441. DOI: 10.1017/CBO9780511804441.

[20] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition", *Data mining and knowledge discovery*, vol. 2, no. 2, pp. 121–167, 1998. DOI: 10.1023/A:1009715923555.

[21] G. Calafiore, "Approximation of n-dimensional data using spherical and ellipsoidal primitives", *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 32, no. 2, pp. 269–278, 2002. DOI: 10.1109/TSMCA.2002.1021114.

[22] E. J. Candes and T. Tao, "Decoding by linear programming", *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, 2005. DOI: 10.1109/TIT.2005.858979.

[23] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The convex geometry of linear inverse problems", *Foundations of Computational mathematics*, vol. 12, no. 6, pp. 805–849, 2012. DOI: 10.1007/s10208-012-9135-7.

[24] J. Chinneck, *Feasibility and Infeasibility in Optimization: Algorithms and Computational Methods*. Cambridge, UK: Springer Science+Business Media, LLC, 2008, ISBN: 9780387749327. DOI: 10.1007/978-0-387-74932-7.

[25] P. L. Combettes and M. R. Civanlar, "The foundations of set theoretic estimation", in *[Proceedings] ICASSP 91: 1991 International Conference on Acoustics, Speech, and Signal Processing*, IEEE, 1991, pp. 2921–2924. DOI: 10.1109/5.214546.

[26] C. Cortes and V. Vapnik, "Support-vector networks", *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995. DOI: 10.1007/BF00994018.

[27] N. R. Council, *Assessing the Reliability of Complex Models: Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification*. Washington, DC: The National Academies, 2012. DOI: 10.17226/13395.

[28] P. S. Craig, M. Goldstein, A. H. Seheult, and J. A. Smith, "Bayes linear strategies for history matching of hydrocarbon reservoirs", in *Bayesian Statistics 5, Proceedings of the Fifth Valencia International Meeting*, J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, Eds., Oxford, UK: Clarendon Press, 1996, pp. 69–95.

[29] ——, "Pressure matching for hydrocarbon reservoirs: A case study in the use of bayes linear strategies for large computer experiments", in *Case Studies in Bayesian Statistics*, C. Gatsonis, J. S. Hodges, R. E. Kass, R. E. McCullock, P. Rossi, and N. D. Singpurwalla, Eds., vol. 3, New York: Springer-Verlag, 1997, pp. 36–93. DOI: 10.1007/978-1-4612-2290-3_2.

[30] F. Dabbene and D. Henrion, "Minimum volume semialgebraic sets for robust estimation", *arXiv preprint arXiv:1210.3183*, 2012.

[31] F. Dabbene, D. Henrion, and C. M. Lagoa, "Simple approximations of semialgebraic sets and their applications to control", *Automatica*, vol. 78, pp. 110–118, 2017. DOI: 10.1016/j.automatica.2016.11.021.

[32] D. Donoho, "For most large underdetermined systems of linear equations the minimal l1-norm solution is also the sparsest solution", *Comm. Pure Appl. Math*, vol. 59, pp. 797–829, 2004. DOI: 10.1002/cpa.20132.

[33] J. Doyle, A. Packard, and K. Zhou, "Review of lfts, lmis, and μ", in *Proceedings of the 30th IEEE Conference on Decision and Control*, 1991, pp. 1227–1232. DOI: 10.1109/CDC.1991.261572.

[34] D. E. Edwards, D. Y. Zubarev, A. Packard, W. A. Lester Jr., and M. Frenklach, "Interval prediction of molecular properties in parametrized quantum chemistry", *Physical review letters*, vol. 112, no. 25, p. 253 003, 2014. DOI: 10.1103/PhysRevLett.112.253003.

[35] R. P. Feeley, "Fighting the curse of dimensionality: A method for model validation and uncertainty propagation for complex simulation models", PhD thesis, University of California, Berkeley, CA, 2008.

[36] R. Feeley, M. Frenklach, M. Onsum, T. Russi, A. Arkin, and A. Packard, "Model discrimination using data collaboration", *J. Phys. Chem. A*, vol. 110, no. 21, pp. 6803–6813, Mar. 2006. DOI: 10.1021/jp056309s.

[37] R. Feeley, P. Seiler, A. Packard, and M. Frenklach, "Consistency of a reaction dataset", *J. Phys. Chem. A*, vol. 108, no. 44, pp. 9573–9583, Oct. 2004. DOI: 10.1021/jp047524w.

[38] E. Fogel and Y.-F. Huang, "On the value of information in system identification—bounded noise case", *Automatica*, vol. 18, no. 2, pp. 229–238, 1982. DOI: 10.1016/0005-1098(82)90110-8.

[39] E. Frazzoli, Z.-H. Mao, J.-H. Oh, and E. Feron, "Resolution of conflicts involving many aircraft via semidefinite programming", *Journal of Guidance, Control, and Dynamics*, vol. 24, no. 1, pp. 79–86, 2001. DOI: 10.2514/2.4678.

[40] M. Frenklach, "Transforming data into knowledge—Process Informatics for combustion chemistry", *Proc. Combust. Inst.*, vol. 31, pp. 125–140, 2007. DOI: 10.1016/j.proci.2006.08.121.

[41] M. Frenklach, A. Packard, G. Garcia-Donato, R. Paulo, and J. Sacks, "Comparison of statistical and deterministic frameworks of uncertainty quantification", *SIAM/ASA J. Uncertainty Quantification*, vol. 4, pp. 875–901, 2016. DOI: 10.1137/15M1019131.

[42] M. Frenklach, A. Packard, and P. Seiler, "Prediction uncertainty from models and data", in *Proc. American Control Conference*, Anchorage, Alaska, New York: IEEE, 2002, pp. 4135–4140. DOI: 10.1109/ACC.2002.1024578.

[43] M. Frenklach, A. Packard, P. Seiler, and R. Feeley, "Collaborative data processing in developing predictive models of complex reaction systems", *Int. J. Chem. Kinet.*, vol. 36, no. 1, pp. 57–66, 2004. DOI: 10.1002/kin.10172.

[44] M. Frenklach, "Modeling", in *Combustion chemistry*, W. C. Gardiner, Ed., Springer, 1984, ch. 7, pp. 423–453, ISBN: 978-1-4684-0186-8. DOI: 10.1007/978-1-4684-0186-8.

[45] M. Frenklach, A. Packard, and R. Feeley, "Optimization of reaction models with solution mapping", *Comprehensive Chemical Kinetics*, vol. 42, pp. 243–291, 2007. DOI: 10.1016/S0069-8040(07)42006-4.

[46] M. Frenklach, H. Wang, and M. J. Rabinowitz, "Optimization and analysis of large chemical kinetic mechanisms using the solution mapping method—combustion of methane", *Progress in Energy and Combustion Science*, vol. 18, no. 1, pp. 47–73, 1992. DOI: 10.1016/0360-1285(92)90032-V.

[47] T. Fujie and M. Kojima, "Semidefinite programming relaxation for nonconvex quadratic programs", *Journal of Global Optimization*, vol. 10, pp. 367–380, 1997. DOI: 10.1023/A:1008282830093.

[48] M. J. Garcia, "Uncertainty quantification for state estimation in power systems", Master's thesis, University of California, Berkeley, CA, 2013.

[49] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian data analysis*. Chapman and Hall/CRC, 2013, ISBN: 9781439840955.

[50] M. Goldstein and N. Huntley, "Bayes linear emulation, history matching, and forecasting for complex computer simulators", *Handbook of Uncertainty Quantification*, pp. 9–32, 2017. DOI: 10.1007/978-3-319-12385-1_14.

[51] M. Goldstein and J. Rougier, "Reified bayesian modelling and inference for physical systems", *Journal of Statistical Planning and Inference*, vol. 139, no. 3, pp. 1221–1239, 2009. DOI: 10.1016/j.jspi.2008.07.019.

[52] M. Grant and S. Boyd, *CVX: Matlab software for disciplined convex programming, version 2.1*, http://cvxr.com/cvx, Mar. 2014.

[53] ——, "Graph implementations for nonsmooth convex programs", in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds., http://stanford.edu/~boyd/graph_dcp.html, Springer-Verlag Limited, 2008, pp. 95–110.

[54] A. Hegde, W. Li, J. Oreluk, A. Packard, and M. Frenklach, "Consistency analysis for massively inconsistent datasets in bound-to-bound data collaboration", *SIAM/ASA J. Uncertainty Quantification*, vol. 6, pp. 429–456, 2018. DOI: 10.1137/16M1110005.

[55] D. Higdon, J. Gattiker, B. Williams, and M. Rightley, "Computer model calibration using high-dimensional output", *J. Amer. Stat. Assoc.*, vol. 103, pp. 570–583, 2008. DOI: 10.1198/016214507000000888.

[56] D. Higdon, M. Kennedy, J. C. Cavendish, J. A. Cafeo, and R. D. Ryne, "Combining field data and computer simulations for calibration and prediction", *SIAM J. Sci. Comput.*, vol. 26, pp. 448–466, 2004. DOI: 10.1137/S10648275034266938.

[57] H. Hjalmarsson, "From experiment design to closed-loop control", *Automatica*, vol. 41, no. 3, pp. 393–438, 2005. DOI: 10.1016/j.automatica.2004.11.021.

[58] R. Horst and P. M. Pardalos, *Handbook of global optimization*. Springer Science & Business Media, 1995, vol. 2, ISBN: 9781475753622. DOI: 10.1007/978-1-4615-2025-2.

[59] S. Iavarone, S. T. Smith, P. Smith, and A. Parente, "Collaborative simulations and experiments for a novel yield model of coal devolatilization in oxy-coal combustion conditions", *Fuel Processing Technology*, vol. 166, pp. 86–95, 2017. DOI: 10.1016/j.fuproc.2017.05.023.

[60] S. Iavarone, J. Oreluk, S. T. Smith, A. Hegde, W. Li, A. Packard, M. Frenklach, P. J. Smith, F. Contino, and A. Parente, "Application of bound-to-bound data collaboration approach for development and uncertainty quantification of a reduced char combustion model", *Fuel*, vol. 232, pp. 769–779, 2018. DOI: 10.1016/j.fuel.2018.05.113.

[61] F. John, "Extremum problems with inequalities as subsidiary conditions", in *Traces and emergence of nonlinear programming*, Springer, 2014, pp. 197–215. DOI: 10.1007/978-3-0348-0439-4_9.

[62] K. Keesman, "Membership-set estimation using random scanning and principal component analysis", *Mathematics and Computers in Simulation*, vol. 32, no. 5-6, pp. 535–543, 1990. DOI: 10.1016/0378-4754(90)90009-8.

[63] M. C. Kennedy and A. O'Hagan, "Bayesian calibration of computer models", *J.R. Statist. Soc. B*, vol. 63, pp. 425–464, 2001. DOI: 10.1111/1467-9868.00294.

[64] M. Kojima and M. Yamashita, "Enclosing ellipsoids and elliptic cylinders of semialgebraic sets and their application to error bounds in polynomial optimization", *Mathematical Programming*, vol. 138, no. 1-2, pp. 333–364, 2013. DOI: 10.1007/s10107-012-0515-1.

[65] A. B. Kurzhanski and P. Varaiya, *Dynamics and Control of Trajectory Tubes: Theory and Computation*. Birkhäuser, 1999, ISBN: 9783319102764. DOI: 10.1007/978-3-319-10277-1.

[66] J. B. Lasserre, "A generalization of löwner-john's ellipsoid theorem", *Mathematical Programming*, vol. 152, no. 1-2, pp. 559–591, 2015. DOI: 10.1007/s10107-014-0798-5.

[67] ——, "Global optimization with polynomials and the problem of moments", *SIAM J. Optim.*, vol. 11, pp. 796–817, 2001. DOI: 10.1137/S1052623400366802.

[68] J. B. Lasserre, "Semidefinite programming vs. lp relaxations for polynomial programming", *Mathematics of operations research*, vol. 27, no. 2, pp. 347–360, 2002. DOI: 10.1287/moor.27.2.347.322.

[69] J.-B. Lasserre and V. Magron, "Optimal data fitting: A moment approach", *SIAM Journal on Optimization*, vol. 28, no. 4, pp. 3127–3144, 2018. DOI: 10.1137/18M1170108.

[70] W. Li, A. Hegde, J. Oreluk, A. Packard, and M. Frenklach, "Uniform sampling of a feasible set", *SIAM/ASA J. Uncertainty Quantification*, In Prep.

[71] W. Li, A. Hegde, J. Oreluk, A. Packard, and M. Frenklach, "Representing model discrepancy in bound-to-bound data collaboration", *submitted to SIAM/ASA J. Uncertainty Quantification*, 2020, arXiv preprint arXiv:1907.00886.

[72] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems", *IEEE Signal Processing Magazine*, vol. 27, 2010. DOI: 10.1109/MSP.2010.936019.

[73] MATLAB, *version 9.6.0.1072779 (R2019a)*. Natick, Massachusetts: The MathWorks Inc., 2019.

[74] MATLAB System Identification Toolbox™, *Modeling an aerodynamic body*, Retrieved April 16, 2020 from https://www.mathworks.com/help/ident/examples/modeling-an-aerodynamic-body.html, The MathWorks, Natick, MA, USA, ver. R2019a.

[75] M. D. McKay, R. J. Beckman, and W. J. Conover, "Comparison of three methods for selecting values of input variables in the analysis of output from a computer code", *Technometrics*, vol. 21, no. 2, pp. 239–245, 1979. DOI: 10.2307/1268522.

[76] I. McKinley T. J.and Vernon, I. Andrianakis, N. McCreesh, J. E. Oakley, R. N. Nsubuga, M. Goldstein, and R. G. White, "Approximate bayesian computation and simulation-based inference for complex stochastic epidemic models", *Statist. Sci.*, vol. 33, pp. 4–18, 2018. DOI: 10.1214/17-STS618.

[77] M. Milanese, J. Norton, H. Piet-Lahanier, and É. Walter, *Bounding approaches to system identification*. Springer-Verlag, 1996. DOI: 10.1007/978-1-4757-9545-5.

[78] B. Möller and M. Beer, "Engineering computation under uncertainty–capabilities of non-traditional models", *Computers & Structures*, vol. 86, no. 10, pp. 1024–1041, 2008. DOI: 10.1016/j.compstruc.2007.05.041.

[79] B. K. Natarajan, "Sparse approximate solutions to linear systems", *SIAM J. Comput.*, no. 24, pp. 227–234, 1995. DOI: 10.1137/S0097539792240406.

[80] J. Nie and J. W. Demmel, "Minimum ellipsoid bounds for solutions of polynomial systems via sum of squares", *Journal of Global Optimization*, vol. 33, no. 4, pp. 511–525, 2005. DOI: 10.1007/s10898-005-2099-2.

[81] B. Ninness and G. Goodwin, "Estimation of model quality", *Automatica*, vol. 31, pp. 1771–1797, 1995. DOI: 10.1016/0005-1098(95)00108-7.

[82] B. M. Ninness and G. C. Goodwin, "Rapprochement between bounded-error and stochastic estimation theory", *International Journal of Adaptive Control and Signal Processing*, vol. 9, no. 1, pp. 107–132, 1995. DOI: 10.1002/acs.4480090111.

[83] J. Nocedal and S. Wright, *Numerical optimization*. Springer Science & Business Media, 2006, ISBN: 9780387400655. DOI: 10.1007/978-0-387-40065-5.

[84] W. Oberkampf and C. Roy, *Verification and Validation in Scientific Computing*. Cambridge, UK: Cambridge University Press, 2010, ISBN: 9780521113601. DOI: 10.1017/CBO9780511760396.

[85] A. O'Hagan, "Bayesian analysis of computer code outputs: A tutorial", *Reliability Engineering & System Safety*, vol. 91, no. 10-11, pp. 1290–1300, 2006. DOI: 10.1016/j.ress.2005.11.025.

[86] J. Oreluk, "Role of experimental data in validating and quantifying uncertainties in complex physical systems", PhD thesis, University of California, Berkeley, CA, 2019.

[87] J. D. Orth, I. Thiele, and B. Ø. Palsson, "What is flux balance analysis?", *Nat. Biotechnol.*, vol. 28, pp. 245–248, 2010. DOI: 10.1038/nbt.1614.

[88] E. Osuna, R. Freund, and F. Girosi, *Support vector machines: Training and applications*, 1997.

[89] B. Ø. Palsson, *Systems Biology: Constraint-based Reconstruction and Analysis*. Cambridge, UK: Cambridge University Press, 2015. DOI: 10.1017/CBO9781139854610.

[90] P. Parrilo, "Semidefinite programming relaxations for semialgebraic problems", *Math. Program., Ser. B*, vol. 96, pp. 293–320, 2003. DOI: 10.1007/s10107-003-0387-5.

[91] J. Pedel, J. N. Thornock, and P. J. Smith, "Ignition of co-axial turbulent diffusion oxy-coal jet flames: Experiments and simulations collaboration", *Combust. Flame*, vol. 160, pp. 1112–1128, 2013. DOI: 10.1016/j.combustflame.2013.01.022.

[92] S. Poljak, F. Rendl, and H. Wolkowicz, "A recipe for semidefinite relaxation for (0, 1)-quadratic programming", *Journal of Global Optimization*, vol. 7, no. 1, pp. 51–73, 1995. DOI: 10.1007/BF01100205.

[93] N. D. Price, J. L. Reed, and B. Ø. Palsson, "Genome-scale models of microbial cells: Evaluating the consequences of constraints", *Nature Reviews Microbiology*, vol. 2, pp. 886–897, 2004. DOI: 10.1038/nrmicro1023.

[94] L. Pronzato and E. Walter, "Minimal volume ellipsoids", *International Journal of Adaptive Control and Signal Processing*, vol. 8, no. 1, pp. 15–30, 1994. DOI: 10.1002/acs.4480080103.

[95] F. Pukelsheim, "The three sigma rule", *The American Statistician*, vol. 48, no. 2, pp. 88–91, 1994. DOI: 10.2307/2684253.

[96] J. B. Rosen, "Pattern separation by convex programming", *Journal of Mathematical Analysis and Applications*, vol. 10, no. 1, pp. 123–134, 1965. DOI: 10.1016/0022-247X(65)90150-2.

[97] T. M. Russi, "Uncertainty quantification with experimental data and complex system models", PhD thesis, University of California, Berkeley, CA, 2010.

[98] T. Russi, A. Packard, R. Feeley, and M. Frenklach, "Sensitivity analysis of uncertainty in model prediction", *J. Phys. Chem. A*, vol. 112, no. 12, pp. 2579–2588, Feb. 2008. DOI: 10.1021/jp076861c.

[99] T. Russi, A. Packard, and M. Frenklach, "Uncertainty quantification: Making predictions of complex reaction systems reliable", *Chem. Phys. Lett.*, vol. 499, pp. 1–8, 2010. DOI: 10.1016/j.cplett.2010.09.009.

[100] J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn, "Design and analysis of computer experiments", *Statistical science*, pp. 409–423, 1989. DOI: 10.1214/ss/1177012413.

[101] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola, *Global sensitivity analysis: the primer*. John Wiley & Sons, 2008, ISBN: 9780470725184. DOI: 10.1002/9780470725184.

[102] P. Seiler, *Nqcqp notes*, Retrieved April 16, 2020, from https://www.csdy.umn.edu/~SeilerControl/Papers/2002/Seiler_02Notes_NonconvexQCQP.pdf.

[103] P. Seiler, M. Frenklach, A. Packard, and R. Feeley, "Numerical approaches for collaborative data processing", *Optim. Eng.*, vol. 7, no. 4, pp. 459–478, Dec. 2006. DOI: 10.1007/s11081-006-0350-4.

[104] H. D. Sherali and W. P. Adams, *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*. Dordrecht: Kluwer Academic Publishers, 1999, ISBN: 978-1-4757-4388-3. DOI: 10.1007/978-1-4757-4388-3.

[105] N. A. Slavinskaya, M. Abbasi, J. H. Starcke, R. Whitside, A. Mirzayeva, U. Riedel, W. Li, J. Oreluk, A. Hegde, A. Packard, M. Frenklach, G. Gerasimov, and O. Shatalov, "Development of an uq-predictive chemical reaction model for syngas combustion", *Energy & Fuels*, vol. 31, no. 3, pp. 2274–2297, 2017. DOI: 10.1021/acs.energyfuels.6b02319.

[106] G. P. Smith, M. Frenklach, R. Feeley, A. Packard, and P. Seiler, "A system analysis approach for atmospheric observations and models: The mesospheric HO_x dilemma", *J. Geophys. Res. (Atmospheres)*, vol. 111, no. D23301, 2006. DOI: 10.1029/2005JD006846.

[107] G. P. Smith, D. M. Golden, M. Frenklach, N. W. Moriarty, B. Eiteneer, M. Goldenberg, C. T. Bowman, R. K. Hanson, S. Song, W. C. Gardiner Jr., V. V. Lissianski, and Z. Qin, *GRI-Mech 3.0*, Retrieved April 16, 2020, from http://combustion.berkeley.edu/gri-mech/.

[108] R. Smith and J. Doyle, "Model validation: A connection between robust control and identification", *IEEE Trans. Automat. Control*, vol. 37, pp. 942–952, 1992. DOI: 10.1109/9.148346.

[109] I. M. Sobol, "On the distribution of points in a cube and the approximate evaluation of integrals", *USSR Computation Mathematics and Mathematical Physics*, vol. 7, no. 4, pp. 86–112, 1967. DOI: 10.1016/0041-5553(67)90144-9.

[110] P. B. Stark, "Constraints versus priors", *SIAM/ASA Journal on Uncertainty Quantification*, vol. 3, no. 1, pp. 586–598, 2015. DOI: 10.1137/130920721.

[111] J. F. Sturm, *Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones*, 1998.

[112] P. Sun and R. M. Freund, "Computation of minimum-volume covering ellipsoids", *Operations Research*, vol. 52, no. 5, pp. 690–706, 2004. DOI: 10.1287/opre.1040.0115.

[113] S. Surjanovic and D. Bingham, *Virtual library of simulation experiments: Test functions and datasets*, Retrieved April 2, 2019, from http://www.sfu.ca/~ssurjano.

[114] M. J. Todd, *Minimum-Volume Ellipsoids: Theory and Algorithms*. Cambridge, UK: Mathematical Optimization Society (MOS), Society of Industrial, and Applied Mathematics (SIAM), 2016, ISBN: 9781611974379. DOI: 10.1137/1.9781611974386.

[115] L. Vandenberghe and S. Boyd, "Semidefinite programming", *SIAM Review*, vol. 38, pp. 49–95, 1996. DOI: 10.1137/1038003.

[116] I. Vernon, M. Goldstein, and R. Bower, "Galaxy formation: A bayesian uncertainty analysis", *Bayesian Analysis*, vol. 5, no. 4, pp. 619–670, 2010. DOI: 10.1214/10-BA524.

[117] ——, "Galaxy formation: Bayesian history matching for the observable universe", *Statistical Science*, vol. 29, no. 1, pp. 81–90, 2014. DOI: 10.1214/12-STS412.

[118] I. Vernon, J. Liu, M. Goldstein, J. Rowe, J. Topping, and K. Lindsey, "Bayesian uncertainty analysis for complex systems biology models: Emulation, global parameter searches and evaluation of gene functions", *BMC Systems Biology*, vol. 12, 2018. DOI: 10.1186/s12918-017-0484-3.

[119] E. Walter and H. Piet-Lahanier, "Estimation of parameter bounds from bounded-error data: A survey", *Mathematics and Computers in simulation*, vol. 32, no. 5-6, pp. 449–468, 1990. DOI: 10.1016/0378-4754(90)90002-Z.

[120] D. Williamson, M. Goldstein, L. Allison, A. Blaker, P. Challenor, L. Jackson, and K. Yamazaki, "History matching for exploring and reducing climate model parameter space using observations and a large perturbed physics ensemble", *Climate dynamics*, vol. 41, no. 7-8, pp. 1703–1729, 2013. DOI: 10.1007/s00382-013-1896-4.

[121] H. Wolkowicz, R. Saigal, and L. Vandenberghe, *Handbook of semidefinite programming: theory, algorithms, and applications*. Springer Science & Business Media, 2000, ISBN: 9781461543817. DOI: 10.1007/978-1-4615-4381-7.

[122] T. M. Yi, M. Fazel, X. Liu, T. Otitoju, J. Goncalves, A. Papachristodolou, S. Prajna, and J. Doyle, "Application of robust model validation using sostools to the study of G-protein signaling in yeast", in *Proceedings of Foundations of System Biology in Engineering*, 2005, pp. 133–136.

[123] X. You, T. Russi, A. Packard, and M. Frenklach, "Optimization of combustion kinetic models on a Feasible Set", *Proc. Combust. Inst.*, vol. 33, pp. 509–516, 2011. DOI: 10.1016/j.proci.2010.05.016.

[124] E. Zamora-Sillero, M. Hafner, A. Ibig, J. Stelling, and A. Wagner, "Efficient characterization of high-dimensional parameter spaces for systems biology", *BMC systems biology*, vol. 5, no. 142, 2011. DOI: 10.1186/1752-0509-5-142.