

Model Adaptivity for Goal-Oriented Inference

by

Harriet Li

B.S. Aerospace Engineering (2013)
Massachusetts Institute of Technology

Submitted to the Department of Aeronautics and Astronautics
in partial fulfillment of the requirements for the degree of

Master of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

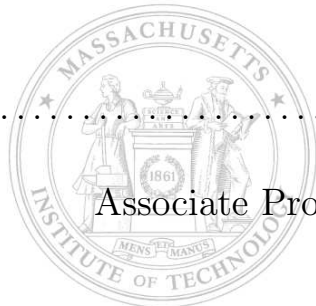
June 2015

© Massachusetts Institute of Technology 2015. All rights reserved.

Author.....
Department of Aeronautics and Astronautics
May 21, 2015

Certified by
Karen Willcox
Professor of Aeronautics and Astronautics
Thesis Supervisor

Accepted by.....
Paulo Lozano
Associate Professor of Aeronautics and Astronautics
Chair, Graduate Student Committee



Model Adaptivity for Goal-Oriented Inference

by

Harriet Li

Submitted to the Department of Aeronautics and Astronautics
on May 21, 2015, in partial fulfillment of the
requirements for the degree of
Master of Science

Abstract

Solving the inverse problem can require many model simulations, which may be expensive for complex models. In the case where the goal of inferring the model parameters is a prediction quantity of interest, we present a method for managing the use of different models in the inference process. The method gives an estimate of the error in the prediction quantity of interest due to the use of a lower fidelity model. Numerical results for the application of the method to 2D models are given.

Thesis Supervisor: Karen Willcox

Title: Professor of Aeronautics and Astronautics

Acknowledgments

...

Contents

1	Introduction	11
1.1	Motivation	11
1.2	Previous Work	12
1.2.1	Goal-Oriented Approaches	12
1.2.2	Mixed-fidelity Models	13
1.3	Thesis Objectives	14
1.4	Thesis Outline	14
2	Mathematical Formulation	17
2.1	Problem Setup	17
2.2	Derivation of Error Estimate	18
2.2.1	Augmented Lagrangian	18
2.2.2	...?	19
2.2.3	Adjoint Formulation	20
2.2.4	Error Estimate	21
2.3	Limitations	22
3	Numerical Results	25
3.1	Convection-Diffusion and Convection-Diffusion-Reaction Models	25
3.1.1	Problem Setup	25
3.1.2	Adaptive Model Refinement	27
3.1.3	Interaction of Observations and QoI	27

3.1.4	...?	32
3.2	Constant versus Field Parameters	33
3.2.1	Problem Setup	33
3.2.2	Adaptive Model Refinement	34
3.3	Cost Analysis	34
4	Conclusion	39
4.1	Thesis Summary	39
4.2	Future Work	40
A	Notation	43
	Bibliography	45

List of Figures

3-1	Locations of observations and QoI region	27
3-2	Element-wise decomposition of error estimate and domain division for mixed-fidelity models	28
3-3	True and estimated absolute relative error in QoI	29
3-4	Effectivity index of QoI error estimate	29
3-5	Element-wise decomposition of error estimates for varying QoI region Ω_I and same three observations	30
3-6	Element-wise decomposition of error estimates for varying observations and same QoI region Ω_I	31
3-7	Estimated absolute relative error in QoI	33
3-8	Element-wise decomposition of error estimate and domain division for mixed-fidelity models	36
3-9	True and estimated absolute relative error in QoI	37
3-10	Effectivity Index of QoI Error Estimate	37

Chapter 1

Introduction

1.1 Motivation

In scientific and engineering contexts, physical systems are represented by mathematical models, which usually take the form of partial differential equations (PDEs). These models are characterized by a set of parameters, and since they relate these parameters to predicted observations, they are referred to as forward models; the forward problem, then, involves solving the equations for a given set of parameters. Inverse problems, on the other hand, refer to when the parameters are unknown and one tries to infer these parameters based on actual observations [4, 17]. Inverse problems arise in many contexts, such as heat transfer [2], medical imaging [3, 15], contaminant source identification [16], and reservoir characterization [13]. In such applications the parameters may be numerous, and yet what is ultimately of interest is some low-dimensional Quantity of Interest (QoI). Thus it may not be necessary to fully resolve all the parameters to obtain the QoI accurately.

In this spirit, one way to relax the process of inferring the parameters is to represent the physical system to a coarser degree. A given physical system can be represented with varying degrees of fidelity by different models. A high-fidelity model may, for example, take into account more physical laws or be more finely discretized, and thus more accurately represent reality. However, a high-fidelity model is also usually

more difficult to solve. Since solving the inverse problems generally requires many evaluations of the forward model, it may be cheaper to use a lower-fidelity model. In addition, the inverse problem is often ill-posed without regularization; it may be the case that even if one were to use the high-fidelity model, the additional resolution (of space-time and/or physical laws) of the high-fidelity model might not even be informed by the observations. To address this, we seek a goal-oriented approach to the inverse problem in the context of multiple models.

In the next section, we describe existing work on goal-oriented approaches and multifidelity modeling.

1.2 Previous Work

1.2.1 Goal-Oriented Approaches

Especially in engineering contexts, the ultimate goal of running a forward simulation or inferring for parameters is to calculate some low-dimensional quantity of interest; the exact states or parameters may not otherwise be of interest. Goal-oriented approaches take advantage of this.

In the context of the forward problem, [5, 14, 19] describe methods for goal-oriented mesh-refinement using adjoints; they derive an a posteriori estimate of the error in an output functional and use this estimate to guide adaptive mesh-refinement. A framework for automated mesh-refinement to calculate a QoI to a prescribed accuracy is described in [20]. The work of [12] also describes adjoint methods for the goal-oriented forward problem, but in the context of concurrent multi-scale models, which are discussed in more detail in Section 1.2.2.

Work has also been done on goal-oriented methods for the inverse problem. Mesh-refinement in the goal-oriented inverse problem¹ is addressed in [6]; they derive an a posteriori estimate of the error in the QoI caused by discretizing the infinite dimen-

¹What we refer to as the goal-oriented inverse problem is referred to as the problem of model calibration in [6].

sional inverse problem, and this error estimate is used to adaptively refine the mesh. For a discretized linear inverse problem, [9] describes a low-dimensional subspace of the parameter space that is both informed by observations and informative to the QoI, and uses it to produce a low-dimensional map from the observations directly to the QoI, sacrificing accuracy in the inferred parameters for accuracy in the QoI that is computed from them.

1.2.2 Mixed-fidelity Models

A particular physical system can be represented with varying degrees of fidelity by different models. Often a lower-fidelity model will be cheaper to solve. However, it will also describe reality to a lesser degree of accuracy; it may include fewer physical laws (for example, a model of Stokes flow ignores viscosity) or describe phenomena at a coarser scale (for example, a model of linear elasticity does not treat individual atoms). A mixed-fidelity model can combine higher- and lower-fidelity models in a way so as to be more tractable than the high-fidelity model, but also have a relatively small error measure.

Two main strategies exist for combining models: hierarchical and concurrent methods. Hierarchical methods (also known as information-passing or sequential methods) take the results of a simulation using the high-fidelity model and use them to inform a lower-fidelity model that is used globally (for example, modeling the molecular structure of a material to determine parameters for constitutive equations). Concurrent methods simultaneously solve the high-fidelity model in some small portions of interest of the domain and the low-fidelity model in the remainder of the domain; for example, [8] apply atomistic models capable of describing bond-breaking behaviors to small clusters of atoms in regions relevant to the formation of fractures, while a continuum model is applied in the rest of the domain.

In our approach, we focus on concurrent methods of combining models. In general, the different models need not depict different scales. For example, [18] form a mixed-fidelity model by dividing the domain into subdomains where either the linear Stokes

equation (low-fidelity) or the nonlinear Navier-Stokes equation (high-fidelity) is solved, based on an a posteriori estimate of the error in a QoI. In addition, the subdomains where the high- and low-fidelity models are solved generally need not be completely disjoint; [1] concurrently solve with atomistic and continuum models which overlap in a “handshake” region, whereas in [18] the subdomains of the high- and low-fidelity models are complementary. The details of how to interface the different models is problem-dependent.

1.3 Thesis Objectives

We aim to combine the ideas in goal-oriented methods for inverse problems, and goal-oriented model adaptivity approaches for forward modeling. The objective of this work is to formulate a method that allows one to systematically manage the use of multiple models in the context of the goal-oriented inverse problem, so as to minimize the error in a QoI prediction.² To do this, we first assume that one can most accurately infer the parameters, and thus most accurately calculate the QoI, by using the highest-fidelity forward model available, but that solving the inverse problem with this model is prohibitively expensive. We derive an estimate for the error in the QoI from inferring the parameters using a lower-fidelity model. This estimate can be localized to individual elements, and the element-wise decomposition can then be used to guide the formation of mixed-fidelity models with which to solve the inverse problem, while minimizing the error in the QoI.

1.4 Thesis Outline

In Chapter 2, we define the goal-oriented inverse problem, derive an a posteriori error estimate for the QoI, and describe its limits and uses for model adaptivity in the context of the goal-oriented inverse problem. In Chapter 3, we apply the error estimate

²Would the little cartoon (showing LF, HF, and MF models with observation and qoi region) from the quals presentation be helpful?

to adaptively form a mixed-fidelity model for models in 2D. In Chapter 4, we give a summary of the thesis, and suggest directions for future work. A summary of all notation used in this thesis can be found in Appendix A.

Chapter 2

Mathematical Formulation

In Section 2.1, we define the goal-oriented inverse problem. In Section 2.2, we derive an a posteriori estimate for the error in the QoI, as compared to that which would have resulted from solving the inverse problem with a high-fidelity model. In Section 2.3, we discuss the limitations of the error estimate.

2.1 Problem Setup

Consider a model described by a PDE, for which the Galerkin formulation of the weak form is written as

$$a(u, \phi; q) = \ell(\phi; q), \quad \forall \phi \in U, \quad (2.1)$$

where $u \in U$ is the state, $q \in Q$ are the unknown parameters, and U, Q are Hilbert spaces.

We define an observation operator $C : U \rightarrow \mathbb{R}^{n_d}$ that maps the state to n_d predicted observations; we denote the actual observations by $d \in \mathbb{R}^{n_d}$. The unknown parameters can then be inferred by minimizing the difference between the predicted and actual observations. This inverse problem is often ill-posed; the observations are noisy and sparse and thus insufficiently informative to uniquely determine the parameters. To remedy this, regularization is used to inject prior information or beliefs about the parameters. The inverse problem with regularization can be written as a constrained

optimization problem

$$\begin{aligned} \text{Minimize} \quad & J(q, u) = \frac{1}{2} \|d - C(u)\|_2^2 + R(q) \\ \text{s.t.} \quad & a(u, \phi; q) = \ell(\phi; q), \quad \forall \phi \in U, \end{aligned} \tag{2.2}$$

where $R(q)$ is the regularization term.

In the case of a goal-oriented inverse problem, the ultimate purpose of inferring the unknown parameters is to calculate some Quantity of Interest (QoI). We denote this QoI by $I(q, u)$, where $I : Q \times U \rightarrow \mathbb{R}$ is a functional that maps the parameters and state to our QoI.

In the following derivation, we assume a and C are three times continuously differentiable, and that I is differentiable.

2.2 Derivation of Error Estimate

A given physical system need not have a unique model a that can describe it; there may be various different models of different fidelities. Assuming that one can most accurately infer the parameters, and thus most accurately calculate the QoI, by using the highest-fidelity forward model available, we derive in this section an a posteriori estimate for the error in the QoI from inferring the parameters with a lower-fidelity model, as compared to that which would have resulted from solving the inverse problem with the highest-fidelity model.

2.2.1 Augmented Lagrangian

The inverse problem can be written as a constrained optimization problem, described in Equation (2.2). Solving this constrained optimization problem is equivalent to finding the stationary point of the corresponding Lagrangian

$$\mathcal{L}(q, u, z) = J(q, u) - (a(u, z; q) - \ell(z; q)), \tag{2.3}$$

where $z \in U$ is the adjoint.

Let $\xi = (q, u, z)$ be called the primary variables. Following the work of [6], we introduce a set of auxiliary variables $\chi = (p, v, y) \in Q \times U \times U$ corresponding to these primary variables, and define an augmented Lagrangian

$$\mathcal{M}((q, u, z), (p, v, y)) = I(q, u) + \mathcal{L}'(q, u, z)(p, v, y). \quad (2.4)$$

Let $\Psi = (\xi, \chi)$ denote the stationary point of \mathcal{M} . Note that

$$\mathcal{M}(\Psi) = I(q, u), \quad (2.5)$$

since taking variations of \mathcal{M} with respect to the auxiliary variables recovers the optimality conditions.

Suppose there are multiple models with which we can infer parameters. For the high-fidelity (HF) model we can define an augmented Lagrangian \mathcal{M}_{HF} , and we denote its stationary point by Ψ_{HF} . Similarly, for the lower-fidelity (LF) model we can define an augmented Lagrangian \mathcal{M}_{LF} , and we denote its stationary point by Ψ_{LF} . Then using Equation (2.5), we can write the error in the QoI from solving the inverse problem with a lower-fidelity model rather than the high-fidelity model as

$$\begin{aligned} I(q_{HF}, u_{HF}) - I(q_{LF}, u_{LF}) = \\ \mathcal{M}_{HF}(\Psi_{HF}) - \mathcal{M}_{HF}(\Psi_{LF}) + \mathcal{M}_{HF}(\Psi_{LF}) - \mathcal{M}_{LF}(\Psi_{LF}). \end{aligned} \quad (2.6)$$

2.2.2 ...?

We rewrite the first two terms $\mathcal{M}_{HF}(\Psi_{HF}) - \mathcal{M}_{HF}(\Psi_{LF})$ by extending the work of [6] to a multi-model context. [6] considers the error in the QoI from solving the inverse problem with a discretized model instead of the infinite-dimensional model.

The discretized model is described by

$$a(u_h, \phi_h; q_h) = \ell(\phi_h; q_h), \quad \forall \phi_h \in U_h \subset U, \quad (2.7)$$

where u_h is the state and $q_h \in Q$ are the unknown parameters, discretized on some mesh with element size h . The infinite-dimensional model is described in Equation (2.7). For these two models they derive the expression

$$I(q, u) - I(q_h, u_h) = \frac{1}{2} \mathcal{M}'(\Psi_h)(\Psi - \Psi_h) + \mathcal{R}(e^3) \quad (2.8)$$

where Ψ is the stationary point of \mathcal{M} as defined in Section 2.2.1 and Ψ_h is its discretized counterpart, and where $e = \Psi - \Psi_h$.

Equation (2.8) holds when the infinite-dimensional model and discretized model are replaced with any higher- and lower-fidelity models, respectively. Thus we have that

$$\mathcal{M}_{HF}(\Psi_{HF}) - \mathcal{M}_{HF}(\Psi_{LF}) = \frac{1}{2} \mathcal{M}'_{HF}(\Psi_{LF})(\Psi_{HF} - \Psi_{LF}) + \mathcal{R}(e^3), \quad (2.9)$$

where $e = \Psi_{HF} - \Psi_{LF}$. In the work of [6], the term $\Psi - \Psi_h$ in Equation (2.8) is estimated using interpolation. In our case, we cannot similarly address the term $\Psi_{HF} - \Psi_{LF}$ in Equation (2.9), since we now have different models instead of different discretizations of the same model.

2.2.3 Adjoint Formulation

We take an adjoint approach to obtain the term $\frac{1}{2} \mathcal{M}'_{HF}(\Psi_{LF})(\Psi_{HF} - \Psi_{LF})$ by viewing it as an error in a linear output. As Ψ_{HF} is a stationary point, it satisfies $\mathcal{M}'_{HF}(\Psi_{HF})(\Phi) = 0$. Let this equation be represented in variational form by

$$A_{HF}(\Psi_{HF}, \Phi) = L(\Phi), \quad \forall \Phi \in (Q \times U \times U)^2, \quad (2.10)$$

and define an output

$$\mathcal{Q}(\Psi_{HF}) = \mathcal{M}'_{HF}(\Psi_{LF})(\Psi_{HF}). \quad (2.11)$$

We can then solve

$$A_{HF}(\Phi, \Lambda) = \mathcal{Q}(\Phi) = \mathcal{M}'_{HF}(\Psi_{LF})(\Phi), \quad \forall \Phi \in (Q \times U \times U)^2 \quad (2.12)$$

for the adjoint Λ , which can be used to obtain

$$\mathcal{M}'_{HF}(\Psi_{LF})(\Psi_{HF} - \Psi_{LF}) = -\mathcal{M}'_{HF}(\Psi_{LF})(\Lambda). \quad (2.13)$$

2.2.4 Error Estimate

Combining Equations (2.9) and (2.13), we obtain a third-order expression for the error in the QoI from inferring the parameters of a lower-fidelity model instead of a high-fidelity model:

$$\begin{aligned} I(q_{HF}, u_{HF}) - I(q_{LF}, u_{LF}) = \\ -\frac{1}{2} \mathcal{M}'_{HF}(\Psi_{LF})(\Lambda) + \mathcal{M}_{HF}(\Psi_{LF}) - \mathcal{M}_{LF}(\Psi_{LF}) + \mathcal{R}(e^3). \end{aligned} \quad (2.14)$$

The error estimate (2.14) does not require that the lower-fidelity model be the lowest-fidelity model. Given a low-fidelity model and a high-fidelity model, an intermediate, mixed-fidelity (MF) model can be formed by using the high-fidelity model in some parts of the domain, and the low-fidelity model in the rest of the domain. Just as error estimates can be used to guide mesh-refinement [5], the error estimate (2.14) can be localized to give elemental contributions and used to guide the division of the domain for a mixed-fidelity model. The error estimate can be calculated again, using the mixed-fidelity model as the lower-fidelity model. This process can be repeated, successively increasing the proportion of the domain in which the high-fidelity model is used, until some threshold is reached. A summary of this approach is presented in Algorithm 1.

Algorithm 1 Adaptively build mixed-fidelity model for low error in QoI

```
1: procedure BUILDMF(HF model, LF model, errTol, maxIter)
2:   Let the model  $\text{MF}_0$  be the LF model applied everywhere in the domain.
3:    $i \leftarrow 0$ 
4:   Solve for stationary point  $\Psi_{\text{MF}_0}$  of augmented Lagrangian  $\mathcal{M}_{\text{MF}_0}$ 
5:   Solve for adjoint  $\Lambda_0$ 
6:   Compute QoI error estimate
      $e_{I,0} = -\frac{1}{2}\mathcal{M}'_{\text{HF}}(\Psi_{\text{MF}_0})(\Lambda) + \mathcal{M}_{\text{HF}}(\Psi_{\text{MF}_0}) - \mathcal{M}_{\text{MF}_0}(\Psi_{\text{MF}_0})$ 
7:   while  $i < \text{maxIter}$  and  $|e_{I,i}/I(q_{\text{MF}_i}, u_{\text{MF}_i})| > \text{errTol}$  do
8:     Localize  $e_{I,i}$  and use this element-wise decomposition to guide formation
       of new mixed-fidelity model  $\text{MF}_{i+1}$ 
9:      $i \leftarrow i + 1$ 
10:    Solve for stationary point  $\Psi_{\text{MF}_i}$  of augmented Lagrangian  $\mathcal{M}_{\text{MF}_i}$ 
11:    Solve for adjoint  $\Lambda_i$ 
12:    Compute QoI error estimate
        $e_{I,i} = -\frac{1}{2}\mathcal{M}'_{\text{HF}}(\Psi_{\text{MF}_i})(\Lambda_i) + \mathcal{M}_{\text{HF}}(\Psi_{\text{MF}_i}) - \mathcal{M}_{\text{MF}_i}(\Psi_{\text{MF}_i})$ 
13:   end while
14: return model  $\text{MF}_i$ 
15: end procedure
```

2.3 Limitations

Although Equation (2.14) is exact, in practice the error estimate is only exact in the case where the model PDEs, observation operator C , and the QoI functional I are linear. In the presence of nonlinearities, the error estimate is approximate, since the now generally nonzero higher order terms $\mathcal{R}(e^3)$ are not calculated, and since the adjoint problem (2.12) is linearized about Ψ_{HF} , but is linearized about Ψ_{LF} in practice.

In motivating this approach, it is assumed that one can most accurately infer the parameters, and thus most accurately calculate the QoI, by using the highest-fidelity forward model available, but that solving the inverse problem with this model is prohibitively expensive. It is also assumed that solving the inverse problem with a mixed-fidelity model, where this highest-fidelity model is only used in a portion of the domain, will be cheaper. There is a cost incurred by using our approach to design such a mixed-fidelity model, however, and it will sometimes be the case that the cost of obtaining this mixed-fidelity model exceeds that of just solving the inverse problem

with the highest-fidelity model directly. Naively, if the auxiliary variables χ have n degrees of freedom, they can be found by solving an $n \times n$ linear system, while the adjoint Λ can be found by solving a $2n \times 2n$ linear system. The cost of solving for the auxiliary variables can be reduced by using a technique described in [6], and the cost of solving for the adjoint Λ can be reduced by preconditioning¹. However, whether it is more efficient to just solve the inverse problem with the high-fidelity model of course depends on the number of iterations needed for such a solve, as well as the number needed to solve the intermediate mixed-fidelity models.

The derived error estimate is applicable to a large class of models. The lower-fidelity model could, for example, be a simplified model including fewer physical phenomena, be a reduced-order model, or have a reduced parameter space. The two models could also correspond to two levels of mesh-refinement, though in this case the method described in [6] would be more efficient, since interpolation could be used to estimate $\Psi_{HF} - \Psi_{LF}$ instead. The derived error estimate is not applicable to all models, however. The two models must have a weak form, so this cannot be applied to, for example, a model of chemical reactions using kinetic Monte Carlo. The weak form, observation operator, and QoI functional must also have the degrees of differentiability noted in Section 2.1. The two models must also have some degree of compatibility, in that the variables of one model must have some meaning in the other; Ψ_{LF} must be in a space admissible to \mathcal{M}'_{HF} .

¹citation needed?

Chapter 3

Numerical Results

In this chapter we demonstrate some results from applying our approach to pairs of steady-state models, implemented using the `libMesh` library. In Section 3.1, we demonstrate the results using two models that differ in the physics modeled. In Section 3.2, the two models considered differ in the space to which the parameter belongs.

3.1 Convection-Diffusion and Convection-Diffusion-Reaction Models

3.1.1 Problem Setup

The high- and low-fidelity models are restricted to a rectangular domain Ω , defined as $\Omega = [0, 5] \times [0, 1]$. The high-fidelity model is a single-species convection-diffusion-reaction equation with a nonlinear reaction term, described by

$$k_d \nabla^2 u - \vec{V} \cdot \nabla u + k_r u^2 = f(q), \quad (3.1)$$

where the state u is the species concentration, $f(q)$ is a forcing field described by the parameters, $k_d = 0.1$ is a diffusion coefficient and $k_r = -42.0$ is a reaction coefficient.

The low-fidelity model

$$k_d \nabla^2 u - \vec{V} \cdot \nabla u + k_r u^2 = f(q) \quad (3.2)$$

differs only in the removal of the reaction term. Both models share a common velocity field, described by $\vec{V}(x, y) = (2y(1 - y), 0)$. To form the mixed-fidelity models, we divide the domain into complementary subdomains Ω_{HF} and Ω_{LF} where the high- and low-fidelity models are solved, respectively. The resulting mixed-fidelity models can be described by

$$k_d \nabla^2 u - \vec{V} \cdot \nabla u + k_r^{MF} u^2 = f(q), \quad (3.3)$$

where k_r^{MF} is a piecewise-constant reaction coefficient

$$k_r^{MF} = \begin{cases} -42.0 & \text{if } x \in \Omega_{HF} \\ 0 & \text{if } x \in \Omega_{LF}. \end{cases} \quad (3.4)$$

The Peclet number is small enough everywhere in the domain to make stabilization unnecessary. Homogeneous Dirichlet boundary conditions are applied on the entire boundary of the domain.

We let the unknown parameters we wish to infer correspond to the forcing field, so that $f(q) = q$. Observations $d = (u(0.35, 0.35), u(1.56, 0.61), u(3.1, 0.5))$ from three points in the domain are artificially generated by running the high-fidelity model on a finer mesh. The QoI we wish to calculate is the integral of the state,

$$I(q, u) = \int_{(x,y) \in \Omega_I} u \, dA, \quad (3.5)$$

over a region $\Omega_I = [0.625, 0.875] \times [0.375, 0.625]$. The locations of the observations and the region Ω_I over which the QoI is calculated is shown in Figure 3-1. Since the inverse problem is ill-posed, we use Tikhonov regularization; the regularization term is $\frac{\beta}{2} \int_{\Omega} \|\nabla f(q)\|_2^2 \, dA$, where $\beta = 10^{-5}$ is a regularization coefficient.

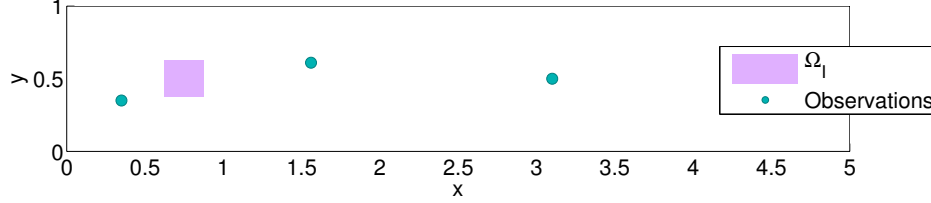


Figure 3-1: Locations of observations and QoI region

3.1.2 Adaptive Model Refinement

The estimated error from Equation (2.14) is then localized to each element to give an element-wise decomposition of the error. Here we calculate the contribution to the estimated error by a particular element κ by simply considering all area integrals in the error estimate expression, limited to that element. As described in Algorithm 1, based on this element-wise decomposition, we increase the proportion of the domain in which the high-fidelity model is used until the estimated absolute relative error in the QoI is less than 1%. Figure 3-2 shows the element-wise decomposition of the error estimate, as well as the subdomains where the low- and high-fidelity models were used, for the series of mixed-fidelity models thus generated. The true and estimated absolute relative errors in the QoI for these same mixed-fidelity models are shown in Figure 3-3, while the effectivity index for the error estimate is shown in Figure 3-4.

It can be seen that in this case, while the error estimates are not exact due to the nonlinear reaction term in the high-fidelity model, the error estimates are fairly accurate. In addition, the QoI that would have been obtained from solving the inverse problem with the high-fidelity model can be replicated to within 1% with a mixed-fidelity model where the high-fidelity model is used in only 20% of the domain.

3.1.3 Interaction of Observations and QoI

The element-wise decomposition of the error estimate (2.14) suggests the use of the high-fidelity model in areas of the domain where the parameter field is both informed by the observations and informative about the QoI. To see this, we compare the element-wise decomposition of the error estimate for three sizes of the QoI region Ω_I

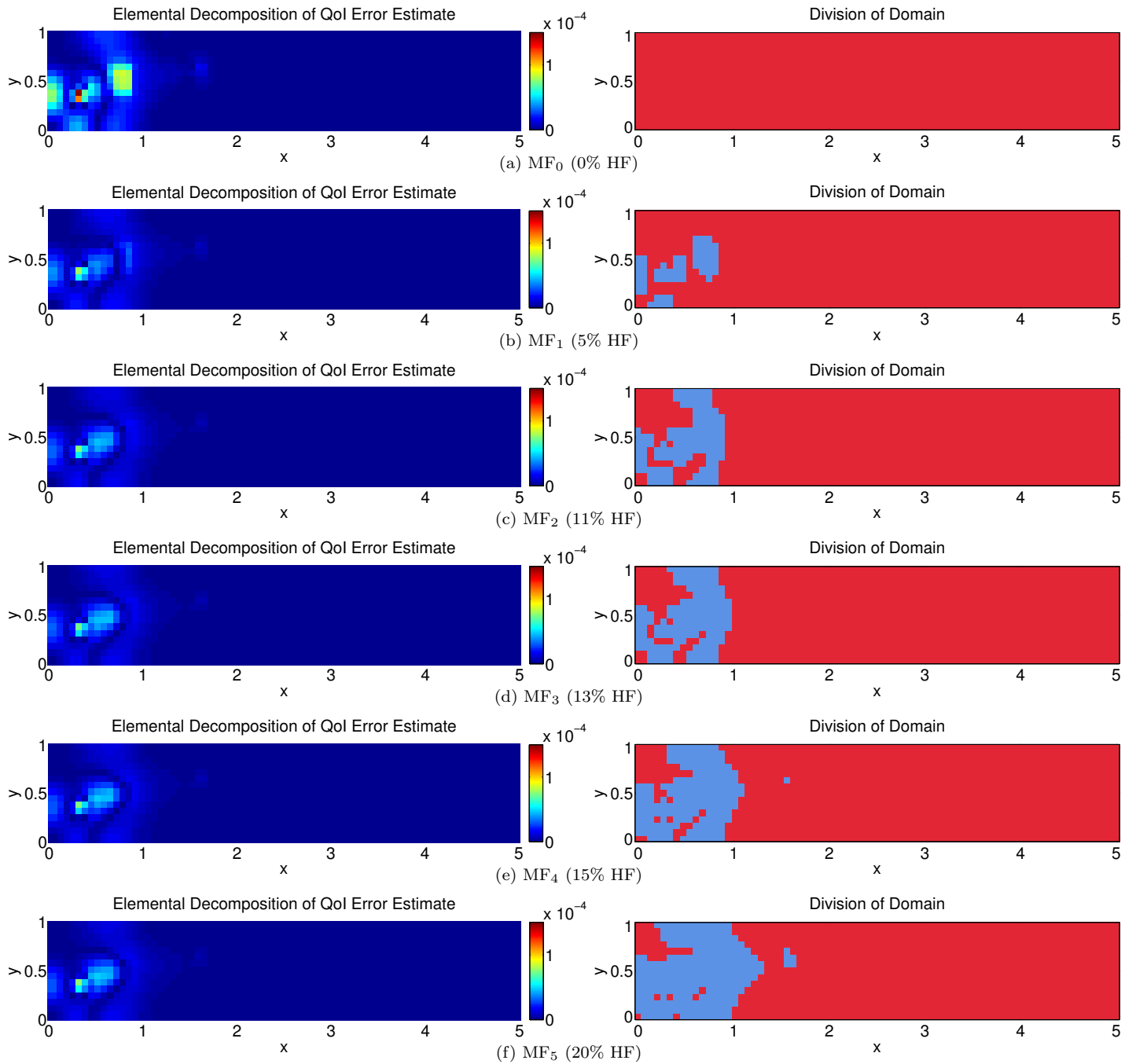


Figure 3-2: Element-wise decomposition of error estimate (left) and domain division (right; low-fidelity model used in red portion, high-fidelity model used in blue portion) for mixed-fidelity models

given the same set of observations, and for three sets of observations given the same QoI region. Again we increase the proportion of the domain in which the high-fidelity

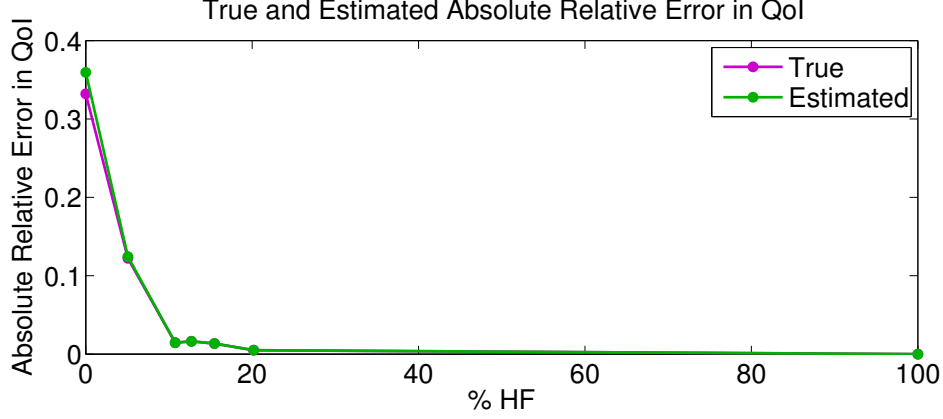


Figure 3-3: True and estimated absolute relative error in QoI

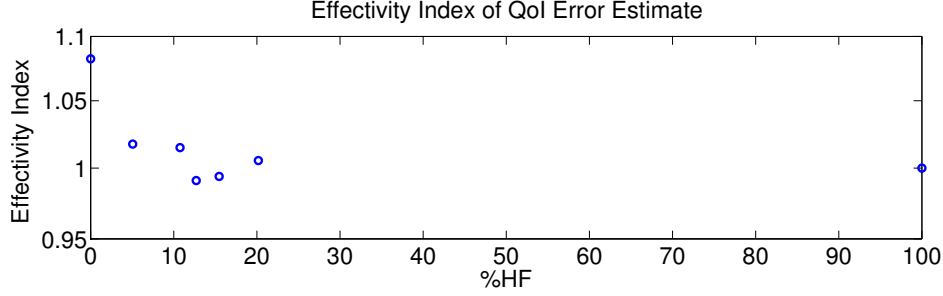


Figure 3-4: Effectivity index of QoI error estimate

model is used until the estimated absolute relative error in the QoI is less than 1%, or until seven iterations have been reached. The element-wise decomposition of the error estimate for the three sizes of the QoI region Ω_I given the same set of observations is shown in Figure 3-5, and for the three sets of observations given the same QoI region is shown in Figure 3-6.

These results make intuitive sense. The regions of the domain where it is most important to use the high-fidelity model are the QoI region, and around those observations made in the QoI region and just upstream of the QoI region; it becomes less important to use the high-fidelity model around observations as they are placed further downstream. In this case, for this pair of models and a QoI that is the integral of the state over a region, an appropriate mixed-fidelity model could have been designed by intuition. The interaction between the observations and the QoI may not always be so intuitive, however, and it is in these cases that a rigorous method for forming a

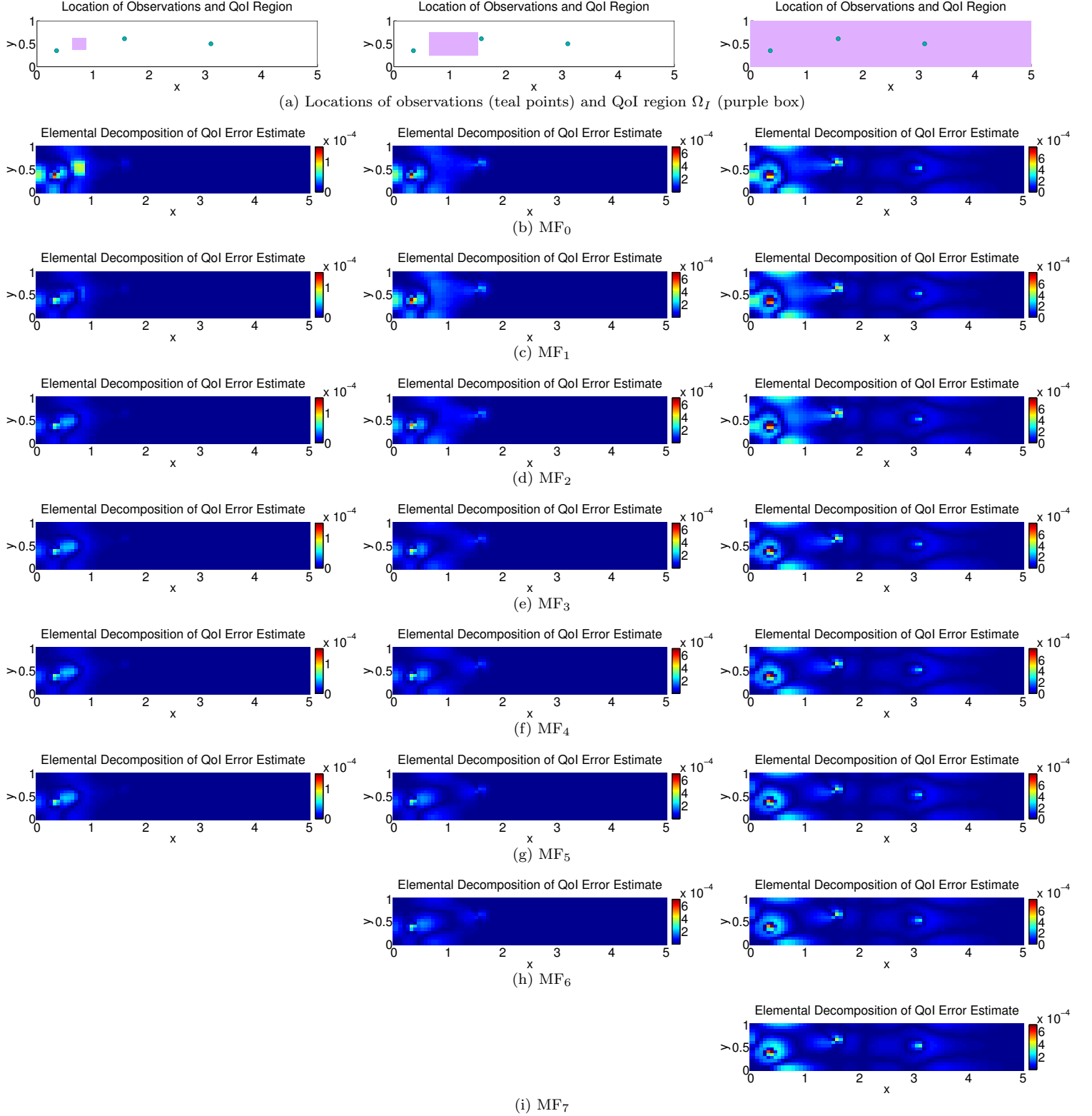


Figure 3-5: Element-wise decomposition of error estimates for varying QoI region Ω_I and same three observations

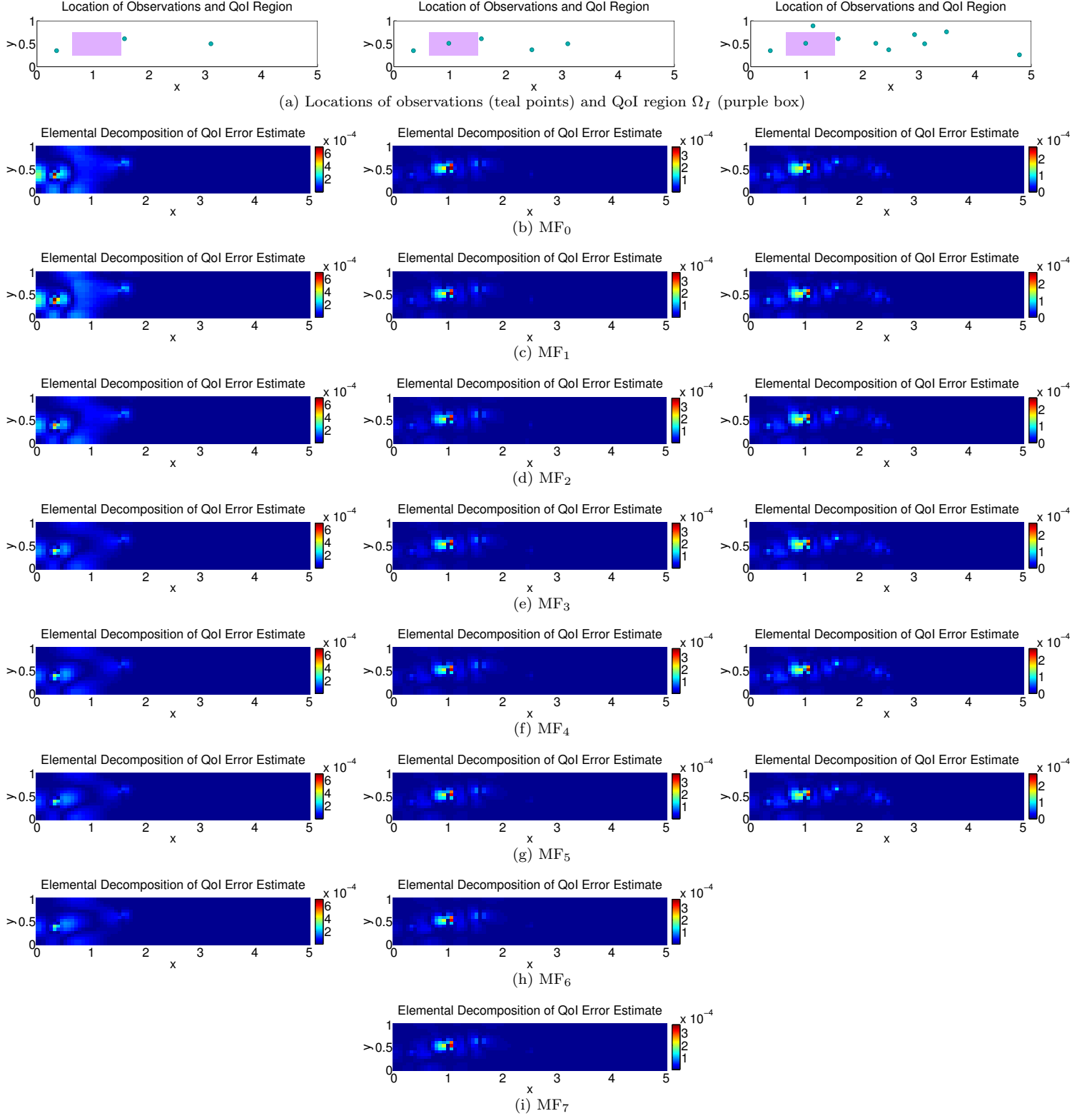


Figure 3-6: Element-wise decomposition of error estimates for varying observations and same QoI region Ω_I

mixed-fidelity model would be most helpful.

3.1.4 ...?

For a highly nonlinear high-fidelity model, it may sometimes be the case that solving the inverse problem with this high-fidelity model requires a specialized optimization algorithm. However, solving the inverse problem with a mixed-fidelity model, where this high-fidelity model is only used in a small portion of the domain, may be achievable using a **generic** optimization algorithm. In this subsection, we give an example of such a case.

To solve the inverse problem, we use the default non-linear solver in **libMesh** (Newton’s method with Brent line-search) to solve the optimality conditions of the corresponding optimization problem (see Equation 2.2). We first solve the inverse problem described in Section 3.1.1 using the high-fidelity model (let it be denoted by HF_{42}) everywhere in the domain, obtaining a solution ξ_{42} . Then, given the same observations, we consider a different high-fidelity model, one where the magnitude of the reaction coefficient in Equation (3.1) is increased from $k_r = -42$ to $k_r = -442$; let this new model be denoted by HF_{442} . We can no longer solve the inverse problem with this new, more non-linear high-fidelity model; using ξ_{42} as an initial guess, the default non-linear solver in **libMesh** fails to converge.

Alternatively, we consider the series of domain divisions for mixed-fidelity models generated using HF_{42} as the high-fidelity model, as described in Section 3.1.2 and shown in Figure 3-2. Using these same domain divisions, we create a series of mixed-fidelity models using the low-fidelity model and the HF_{442} model. The inverse problem can be solved for these mixed-fidelity models using the default non-linear solver, and the estimated relative error in the QoI can be reduced to less than 5%, as shown in Figure 3-7.

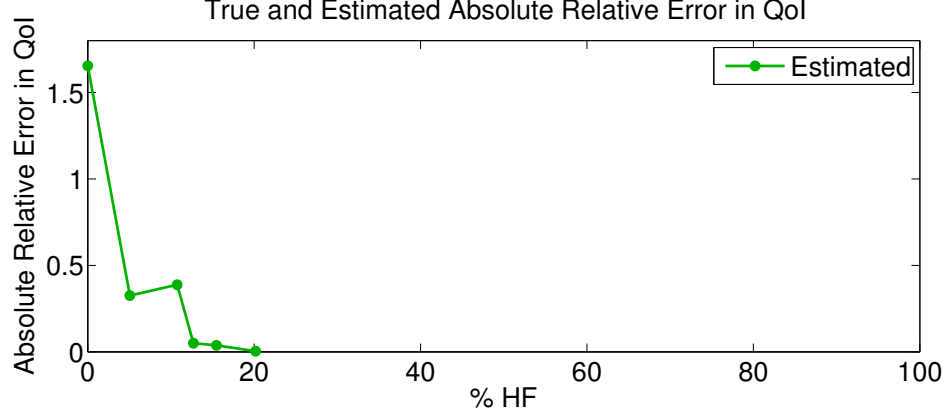


Figure 3-7: Estimated absolute relative error in QoI

3.2 Constant versus Field Parameters

3.2.1 Problem Setup

The high-fidelity model

$$k_d \nabla^2 u - \vec{V} \cdot \nabla u + k_r u^2 = f(q), \quad q \in U, \quad (3.6)$$

is again a single-species convection-diffusion-reaction equation with a nonlinear reaction term, where $k_d = 0.1$ is a diffusion coefficient and $k_r = -4.2$ is a reaction coefficient. The low-fidelity model

$$k_d \nabla^2 u - \vec{V} \cdot \nabla u + k_r u^2 = f(q), \quad q \in \mathbb{R} \quad (3.7)$$

differs from the high-fidelity model only in that the parameter is a constant instead of a field. Then the intermediate mixed-fidelity models have parameter fields which are allowed to be non-constant in only portions of the domain. For ease of implementation, we require that the resulting parameter field remain continuous at the interface between the low-fidelity and high-fidelity subdomains, although this constraint is not necessary for the theory to hold. The velocity field and boundary conditions, as well as the observations, unknown parameters to be inferred, and QoI, remain the same

as described in Section 3.1. As the inverse problem is ill-posed, except for perhaps in the case where the low-fidelity model is used throughout the domain, regularization is added; the Tikhonov regularization term is $\frac{\beta}{2} \int_{\Omega} \|\nabla f(q)\|_2^2 + f(q)^2 \, dA$, where $\beta = 10^{-3}$ is a regularization coefficient.

3.2.2 Adaptive Model Refinement

Based on the element-wise decomposition of the estimated error, we increase the proportion of the domain in which the high-fidelity model is used until the estimated absolute relative error in the QoI is less than 5%. In this case, we allow an element assigned to be part of the high-fidelity portion of the domain in one mixed-fidelity model to be reassigned back to be part of the low-fidelity portion of the domain in subsequent mixed-fidelity models if its contribution to the error is not large enough. Figure 3-8 shows the element-wise decomposition of the error estimate, as well as the subdomains where the low- and high-fidelity models were used, for the first six of the series of mixed-fidelity models thus generated. The true and estimated absolute relative errors in the QoI for these same intermediate models are shown in Figure 3-9, while the effectivity index of the error estimate is shown in Figure 3-10.

As in the previous example, the error estimates are only approximate due to the nonlinear term in both the low- and high-fidelity models. In this case, the high-fidelity model must also be used in a larger portion of the domain (60%) before the estimated relative error in the QoI reaches the desired level.

3.3 Cost Analysis

In both the examples discussed in Sections 3.1 and 3.2, the high-fidelity model is simple and solving the inverse problem with the high-fidelity model is easily achievable. Doing so actually requires less computational time than using Algorithm 1 to rigorously form a mixed-fidelity model with which to solve the inverse problem instead. However, as discussed in Section 2.3, we assume in motivating our approach that solving the inverse

problem with the high-fidelity model is prohibitively expensive; although there is no benefit, in terms of computational cost, to using our approach in the given examples, one would likely see such benefits for more complex models. Such a benefit is suggested in Section 3.1.4, where the inverse problem using the high-fidelity model is difficult to solve but our approach allows one to construct a mixed-fidelity model for which the inverse problem can be solved with a **generic** algorithm and yet produce a QoI with a small estimated relative error.

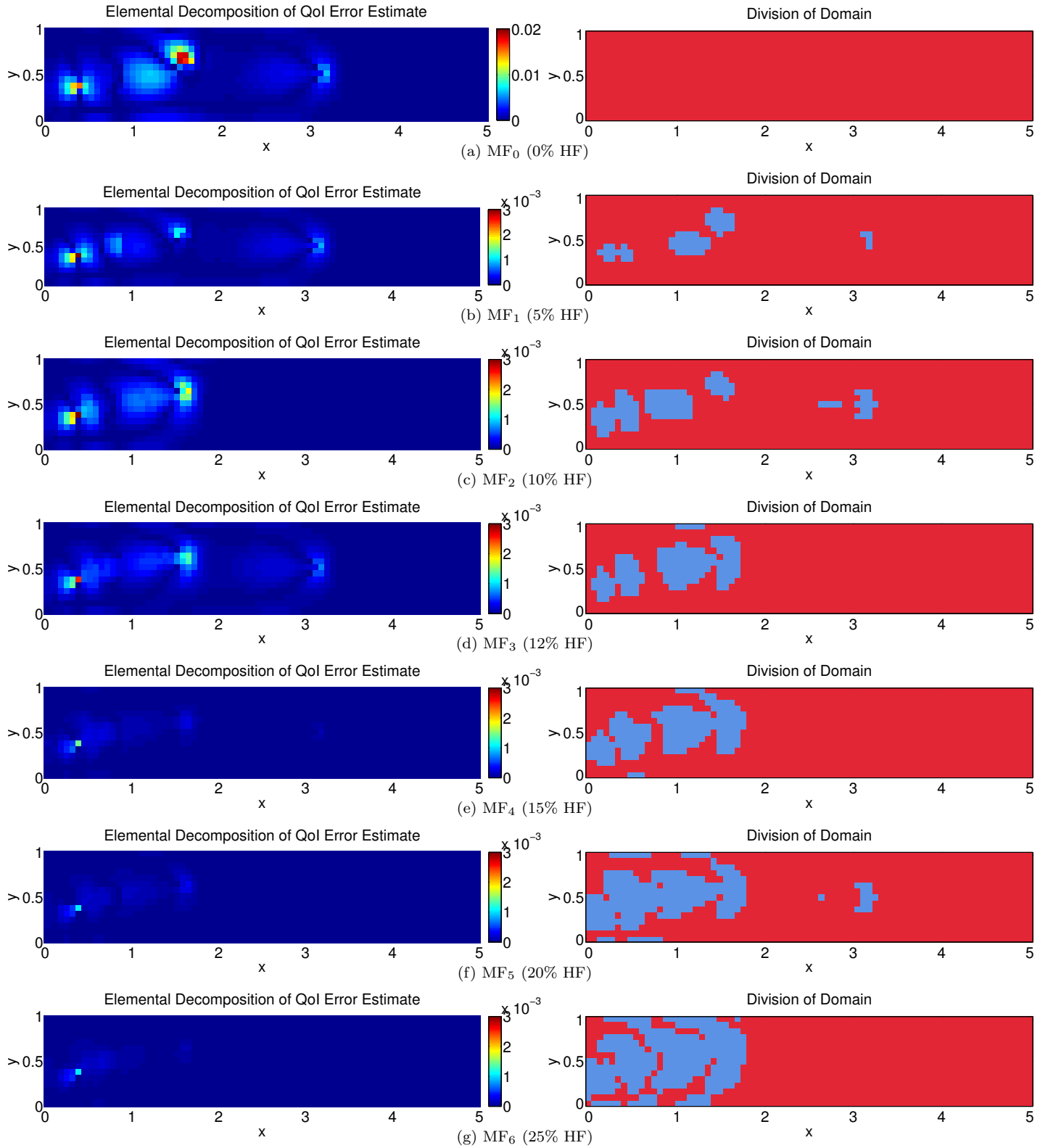


Figure 3-8: Element-wise decomposition of error estimate (left) and domain division (right; low-fidelity model used in red portion, high-fidelity model used in blue portion) for mixed-fidelity models

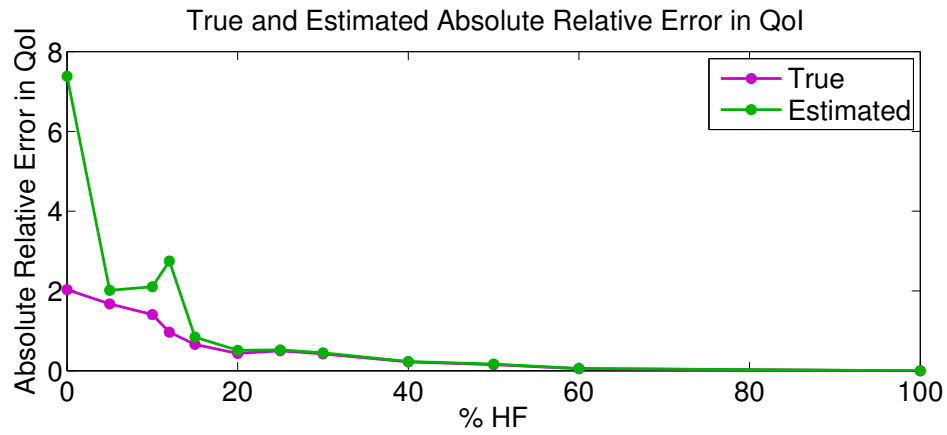


Figure 3-9: True and estimated absolute relative error in QoI

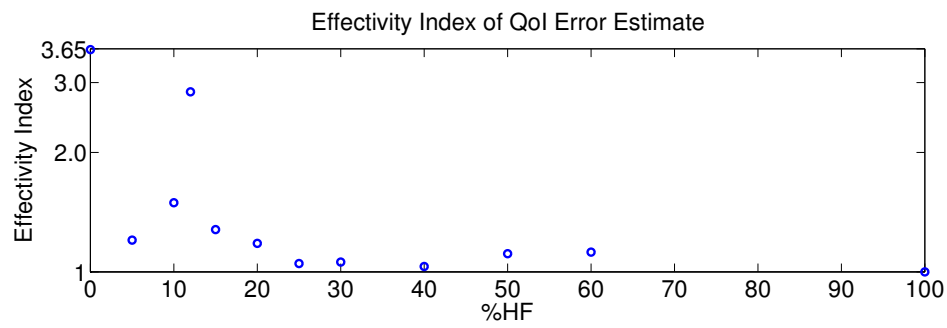


Figure 3-10: Effectivity Index of QoI Error Estimate

Chapter 4

Conclusion

4.1 Thesis Summary

In this work we have presented an error estimator that can be used to adaptively create a mixed-fidelity model with which to solve the inverse problem, so as to minimize the error in the QoI calculated from the inferred parameters. We applied this method to pairs of 2D models, one that differed in the physics included and one that differed in the space to which the parameters belonged; we were able to obtain a value for the QoI with a small relative error without having to solve the inverse problem with the high-fidelity model. In these cases, the element-wise decomposition of the error estimate also indicated regions of the parameter field that were both informed by the observations and informative to the QoI.

The inverse problem with the high-fidelity models examined were not so expensive to solve as to warrant the adaptive formation of a mixed-fidelity model; however, based on existing work with concurrent multi-fidelity models, we are confident that savings in computational cost can be obtained in cases where the high-fidelity model is more complex and solving the inverse problem with the high-fidelity model is less tractable. We demonstrated a case where the inverse problem with the high-fidelity model could not be solved without a specialized nonlinear solver, but where our method resulted in a mixed-fidelity model for which the inverse problem could be solved with a generic

nonlinear solver, with a small relative error in the QoI.

4.2 Future Work

An immediate direction for extension of this work is to the case of the statistical inverse problem. Thus far in this work, we have considered the deterministic inverse problem, as described in Section 2.1; we seek to infer the parameter values that optimally fit the observations and the prior beliefs embedded in the regularization. However, we can rarely, if ever, be certain that the inferred values are correct, whether this be due to epistemic uncertainty from a lack of knowledge or aleatoric uncertainty from inherent variability in the physical system, or both [11]. One may attempt to capture the uncertainty in the inferred parameters by representing them as random variables with a probability distribution; inferring the distribution of the parameters given some observations is the statistical inverse problem.

A popular approach to solving the statistical inverse problem is to apply a Bayesian framework; Bayes' rule is used to combine a prior distribution, which captures prior beliefs about the parameters, and a likelihood distribution, which captures the likelihood of observations given an instance of the parameter values and a model of noise in the observations, to give a posterior distribution on the parameters. Since there is generally no analytical expression for this posterior distribution, it is usually characterized by samples from the distribution. Sampling methods like the popular Markov chain Monte Carlo (MCMC) method require many evaluations of the forward model, and for a large parameter space, this quickly becomes intractable. In engineering contexts, it is still usually the case, however, that we are ultimately interested in a low-dimensional QoI, and it is the uncertainty in this low-dimensional quantity that we wish to capture; this low-dimensional distribution is referred to as the predictive posterior.

One way we could potentially apply this work to the statistical inverse problem is by reducing the parameter space that needs to be sampled. Such a direction is suggested

by the results presented in Section 3.2, where the mixed-fidelity model had significantly fewer degrees of freedom in its parameter field than the high-fidelity model, and thus a smaller parameter space. In the case of a linear model and observation operator, and a Gaussian prior and additive Gaussian noise in the observations, the posterior distribution of the parameters is also Gaussian. There are parallels between the objective function of the deterministic inverse problem with Tikhonov regularization and the mode of the posterior distribution [10], which could potentially be drawn upon to extend this work to the creation of an alternative statistical inverse problem that, by utilizing a mixed-fidelity model with fewer degrees of freedom in its parameter field, requires exploration of a small parameter space with minimal compromise in the predictive posterior.

Another potential approach would be to extend our method to the creation of mixed-fidelity models that are used as surrogates; these surrogate models can be evaluated in place of the high-fidelity model, thus decoupling the number of expensive forward evaluations of the high-fidelity model needed from the number of posterior parameter distribution samples that is desired [7]. The samples obtained using such a surrogate might sacrifice accuracy in representing the posterior parameter distribution for accuracy in representing the predictive posterior distribution.

Appendix A

Notation

For quick reference, Tables A.1 and A.2 summarize notation used throughout this thesis (in alphabetical order).

Table A.1: Summary of Notation (Part 1)

Symbol	Meaning
a	how to describe? not necessarily bilinear...semilinear?
A	similar to above...also how to refer to this adjoint problem?
\mathcal{C}	Observation operator
d	Observations
e	Error in stationary point of augmented Lagrangian
f	Forcing function
HF	High-fidelity
I	QoI functional
J	Objective function of optimization formulation of inverse problem
k_d	Diffusion coefficient
k_r	Reaction coefficient
ℓ	Linear functional; part of weak form of forward model
\mathcal{L}	Lagrangian
L	Linear functional; part of weak form of what?
LF	Low-fidelity
\mathcal{M}	Augmented Lagrangian
MF	Mixed-fidelity
n_d	Number of observations

Table A.2: Summary of Notation (Part 2)

Symbol	Meaning
p	Auxiliary variable corresponding to parameters
ϕ	Test function
q	Unknown parameters
Q	Hilbert space
\mathcal{Q}	Output functional
R	Regularization term
\mathcal{R}	Higher-order terms
u	State variables
U	Hilbert space
v	Auxiliary variable corresponding to state
\vec{V}	Velocity field
y	Auxiliary variables corresponding to adjoint
z	Adjoint variables
β	Regularization coefficient
Γ	Super-adjoint
ξ	Primary variables
χ	Auxiliary variables
Ψ	Stationary point of augmented Lagrangian
Ω	Domain
Ω_{HF}	Subdomain over which high-fidelity model is solved
Ω_I	Region in domain over which QoI is calculated
Ω_{LF}	Subdomain over which low-fidelity model is solved

Bibliography

- [1] Farid F. Abraham, Jeremy Q. Broughton, Noam Bernstein, and Efthimios Kaxiras. Spanning the length scales in dynamic simulation. *Computers in Physics*, 12(6):538–546, 1998.
- [2] Oleg M. Alifanov. *Inverse Heat Transfer Problems*. Springer-Verlag Berlin Heidelberg, 1994.
- [3] S.R. Arridge. Optical tomography in medical imaging. *Inverse Problems*, 15:R41–R93, 1999.
- [4] H.T. Banks and K. Kunisch. *Estimation Techniques for Distributed Parameter Systems*. Birkhäuser, 1989.
- [5] Roland Becker and Rolf Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numerica*, 10:1–102, 2001.
- [6] Roland Becker and Boris Vexler. Mesh refinement and numerical sensitivity analysis for parameter calibration of partial differential equations. *Journal of Computational Physics*, 206:95–110, 2005.
- [7] Patrick R. Conrad, Youseff M. Marzouk, Natesh S. Pillai, and Aaron Smith. Accelerating Asymptotically Exact MCMC for Computationally Intensive Models via Local Approximations. *ArXiv e-prints*, 2014.
- [8] Roopam Khare, Steven L. Mielke, George C. Schatz, , and Ted Belytschko. Multiscale coupling schemes spanning the quantum mechanical, atomistic forcefield, and continuum regimes. *Computer Methods in Applied Mechanics and Engineering*, 197:3190–3202, 2008.
- [9] Chad Lieberman and Karen Willcox. Goal-oriented inference: Approach, linear theory, and application to advection diffusion. *SIAM Review*, 55(3):493–519, 2013.
- [10] James Martin, Lucas C. Wilcox, Carsten Burstedde, and Omar Ghattas. A stochastic newton mcmc method for large-scale statistical inverse problems with application to seismic inversion. *SIAM Journal of Scientific Computing*, 34(3):A1460–A1487, 2012.

- [11] William L. Oberkampf et al. Challenge problems: uncertainty in system response given uncertain parameters. *Reliability Engineering & System Safety*, 85:11–19, 2004.
- [12] J. Tinsley Oden, Serge Prudhomme, Albert Romkes, and Paul T. Bauman. Multi-scale modeling of physical phenomena: Adaptive control of models. *SIAM Journal of Scientific Computing*, 28(6):2359–2389, 2006.
- [13] Dean S. Oliver, Albert C. Reynolds, and Liu Ning. *Inverse Theory for Petroleum Reservoir Characterization and History Matching*. Cambridge University Press, 2008.
- [14] Serge Prudhomme and J. Tinsley Oden. On goal-oriented error estimation for elliptic problems: application to the control of pointwise errors. *Computer Methods in Applied Mechanics and Engineering*, 173:313–331, 1999.
- [15] Kui Ren. Recent developments in numerical techniques for transport-based medical imaging methods. *Communications in Computational Physics*, 8(1):1–50, 2010.
- [16] Ne-Zheng Sun. *Inverse Problems in Groundwater Modeling*, volume 6 of *Theory and Applications of Transport in Porous Media*. Springer Netherlands, 1999.
- [17] Albert Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics, 2005.
- [18] T.M. van Opstal, P.T. Bauman, S. Prudhomme, and E.H. van Brummelen. Goal-oriented model adaptivity for viscous incompressible flows. *Computational Mechanics*, pages 1–10, 2015.
- [19] David A. Venditti and David L. Darmofal. Adjoint error estimation and grid adaptation for functional outputs: Application to quasi-one-dimensional flow. *Journal of Computational Physics*, 164:204–227, 2000.
- [20] Masayuki Yano. *An Optimization Framework for Adaptive Higher-Order Discretizations of Partial Differential Equations on Anisotropic Simplex Meshes*. PhD thesis, Massachusetts Institute of Technology, 2012.