A Bayes Linear Approach to Making Inferences from X-rays

Benjamin Thomas Lopez

A Thesis presented for the degree of

Doctor of Philosophy

Statistics and Probability
Department of Mathematical Sciences
University of Durham
England

September 2018

Dedicated to

Ellerie.

A Bayes Linear Approach to Making Inferences from X-rays

Benjamin Thomas Lopez

Submitted for the degree of Doctor of Philosophy

September 2018

Abstract

X-ray images are often used to make inferences about physical phenomena and the entities about which inferences are made are complex. The Bayes linear approach is a generalisation of subjective Bayesian analysis suited to uncertainty quantification for complex systems. Therefore, Bayes linear is an appropriate tool for making inferences from X-ray images.

In this thesis, I will propose methodology for making inferences about quantities, which may be organised as multivariate random fields. A number of problems will be addressed: anomaly detection, emulation, inverse problem solving and transferable databases. Anomaly detection is deciding whether a new observation belongs to the same population as a reference population, emulation is the task of building a statistical model of a complex computer model, inverse problem solving is the task of making inferences about system values, given an observation of system behaviour and transferable databases is the task of using a data-set created using a simulator to make inferences about physical phenomena.

The methods we use to address these problems will be exemplified using applications from the X-ray industry. Anomaly detection will be used to identify plastic contaminants in chocolate bars, emulation will be used to efficiently predict the scatter present in an X-ray image, inverse problem solving will be used to infer an entity's composition from an X-ray image and transferable databases will be used to improve image quality and return diagnostic measures from clinical X-ray images. The Bayes linear approach to making inferences from an X-ray image enables improvements over the state-of-the-art approaches to high impact problems.

Declaration

The work in this thesis is based on research carried out at the Statistics and Probability Group, the Department of Mathematical Sciences, England. No part of this thesis has been submitted elsewhere for any other degree or qualification and it is all my own work unless referenced to the contrary in the text.

Acknowledgements

Special thanks go to Camila, Michael and Gary for their patient supervision and invaluable incites. Thanks go to Paul and Adam for endless sound-boarding, belief in the methods and knowledge of X-ray physics; Josh for his help with signal processing and Will for his coding expertise. Thanks go to Ibex Innovations for providing interesting problems, resources and data. I would also like to thank Marcelo and Sam for keeping me company in the office and my family for raising and supporting me. Finally, I would thank Ellerie for putting up with me, particularly while writing this thesis!

Contents

List of Figures

Chapter 1

Introduction

At the beginning of this research project, there were a concise set of problem statements. These problems were provided by a company specialising in X-ray examination technology – Ibex Innovations. The problems were chosen because a satisfactory solution would have a high impact on the industry. As with many practical applications, the problems addressed are a special case of a more general class of problems. The aim of this thesis is to set out the general class of problems to which the industrial applications belong, a methodology that may be utilised to solve this class and a solution to the actual applications.

1.1 Problem Statements

1.1.1 Contaminant Detection

Problem Statement

The first problem that was proposed was detecting contamination in edible products. The problem of particular interest was finding plastic contaminants in chocolate bars. This leads to the first problem statement.

Problem Statement 1.1.1. *Given an X-ray image of a chocolate bar, detect plastic contaminants.*

Challenges

Chocolate bars have similar density and chemical composition to plastic, the X-ray images have a large number of elements and inferential calculations must not impede the production process. Therefore, i) the signal from contamination is small and hard to distinguish

from product variation, ii) the computation is complicated by large data-sets and iii) the algorithm must not slow down the production process.

Methods

Our solution to this problem is presented in Chapter 3. The method uses sets of X-ray images of chocolate bars measured with a novel X-ray examination technology. The images are modelled as multivariate random fields. The population to which the images of non-contaminated chocolate bars belong is modelled using second order exchangeability. Whether plastic contamination is present is decided using the difference of a bar from the population given that it is non-contaminated. The results show capability to detect plastic contaminants using X-ray examination for the first time.

In developing a solution to this problem, a number of novel approaches were developed: i) a novel method is developed, named continuous energy, for extracting multi-spectral information from a single radiograph, ii) a Bayes linear anomaly detection method is presented which enables product damage and product contamination to be differentiated and iii) approximate kernel methods enable efficient calculation of adjusted beliefs. Results are presented that show the method presented enable the location of (1-5mm) plastic contaminates to be identified in chocolate bars.

1.1.2 Scatter Removal

The second problem that we address is reducing the effect of photons that have undergone a scattering interaction in clinical X-ray imaging. Scattered photons degrade image quality to such an extent that a clinician's ability to make a diagnosis is impeded. This leads to the second problem statement.

Problem Statement 1.1.2. *Given the chemical composition of an entity being examined, infer the effect of scattered photons on an X-ray image.*

Challenges

The rate at which scattering interactions occur is dependent on the chemical composition of the entity. This poses two challenges: encoding the chemical composition of a complex entity mathematically and describing the complex relationship between an entity's composition and scattering interactions. Mathematical physics models are powerful tools for understanding complex real world phenomena. However, they are often expensive to

evaluate and they cannot model the complex entities of interest. Therefore, this problem is challenging because i) the entities considered are complex so require a large number of parameters to encode their chemical composition, ii) the physics models are under-resolved so the scattering behaviour for entities of interest cannot be revealed and iii) the models are complex so viewing scattering behaviour, even for simple entities, is computationally expensive.

Methods

In Chapter 4, we combine a stochastic physics model of X-ray examination with Bayesian emulation to be able to perform accurate prediction of the intensity due to scattered photons in an X-ray image. The approach uses i) Bayesian emulation to remove the computational impediment and ii) a combination of Bayesian emulation and signal processing approaches to make inferences about scattering behaviour for complex entities using only simulations of entities that are defined with a small (under 6) number of parameters. This enables the assumptions made in current state-of-the-art methods (see Section 4.3.1) to be relaxed, improving the accuracy of scatter prediction. The algorithm is validated in simulation on phantoms designed to be analogous to anatomy commonly examined in clinical practice.

Although emulation has been used to model many simulators, it has not yet been used in scatter estimation. The method presented removes the usual assumption that the point-spread-function (PSFs) are piece-wise-stationary and isotropic. This enables a more accurate model of simulations of X-ray images of entities with complex geometry and, therefore, more accurate scatter estimates.

1.1.3 Material Inference

Problem Statement

Making inferences about a patient's composition from X-ray images is useful to clinicians. For example, areal Bone Mineral Density (aBMD) can be inferred from X-ray images and is useful in diagnosing osteoporosis. Furthermore, as a patient's composition is uncertain, a method for making inferences about their composition is essential to a scatter correction method. This leads to the third problem statement.

Problem Statement 1.1.3. *Given an X-ray image of an entity, infer its material composition.*

Challenges

The problem of inferring an entity's composition is an inverse problem. A simple description of how inverse problems may be solved is to check every possible solution and decide on the set which is an acceptable solution. This approach is impractical when i) the calculation of whether a possible solution is acceptable is expensive and ii) the space of possible solutions is large. Then the number of solutions which need to be checked to effectively explore the solution space is huge and each possible solution takes a long time to check. In this example, the entities of interest are complex − meaning a large number of parameters are required to describe them mathematically − and the X-ray image produced from a given entity is expensive to simulate. Therefore, the space of possible solutions is high dimensional and the cost of checking a proposal solution is high.

Methods

In Chapter 5, we use a Bayesian inverse problem solving approach with multivariate random field priors to solve this problem using mono-spectral X-ray examination equipment. Multivariate random field priors reduce the effective dimension of the solution space and iterative history matching reduces the number of solutions which have to be checked. Results in simulation show that the problem of composition inference is not, in the Bayesian sense, ill posed. This means that composition measures, like aBMD, can be made available from standard digital radiography equipment and scatter removal can be performed even if the entity composition is uncertain.

The fact that composition inference is not ill-posed was a surprising and novel result of this work. The physicists involved in this work were convinced at the outset that a multi-spectral X-ray image would be required. This result means that Bayesian inverse problem solving may be used to create a scatter estimate which depends on the estimated thickness and alloy (not just thickness) of an entity being examined. Furthermore, diagnostics measures (which currently require referral for a specialised multi-spectral radiograph) can be made available from a mono-spectral radiograph. Due to the dimensionality of the input space, two new Bayesian inverse problem solving tools were developed: i) history matching with random field priors and ii) iterative history matching. These approaches

enable an acceptable solution to be found in a useful time-frame.

1.1.4 Transferable Database

Problem Statement

Making inferences for material composition or the scatter present in an X-ray image requires an appropriate training-set. There are three prominent limiting factors when creating appropriate training-sets: resource, ethical and practical. For example, a resource limitation would be when the number of observations required to train a model is too expensive to measure, an ethical limitation would be when the material required for training is un-ethical to attain and a practical limitation would be when the experiment required is impractical to perform in the real world. A solution to this problem is to use a physics model, instead of real world observations, to make inferences. This leads to the final problem statement.

Problem Statement 1.1.4. *Given a simulated database and an X-ray image of an entity, infer its chemical composition and remove the effect of scattered photons.*

Challenges

The physics model we use is a mathematical description of how an entities chemical composition affects the X-ray image produced. However, this relationship is imperfect and there will be discrepancy between the model and the system it attempts to mimic. This imperfection complicates the transferable database problem as the relationship between a simulated database and real world behaviour is unknown.

Methods

In Chapter 6, a subjective Bayesian approach to transferring inferences from complex computer models to physical systems is utilised. This involves a description of the discrepancy between a model and reality, and how observations on similar physical systems may be used to resolve this discrepancy. Results are presented that show real world composition inference and scatter removal on quantitative phantoms and cadavers.

Explicitly considering the difference between a simulator and reality (model discrepancy) in a scatter removal method is a novel output of this work. In all the literature we reviewed that used a simulator to performed scatter removal, not one explicitly included in

their inference the fact that their simulator was not reality. Therefore, making inferences about the difference between the simulator and the system it purports to mimic is a new idea in scatter estimation. In most other applications which consider model discrepancy, the focus is on the single system case. In this work, an extension is developed and observations of multiple systems are used to learn about model discrepancy. This approach enables scatter removal (producing improved image quality over current state-of-the-art methods) and material inference from real world X-ray images.

In the process of creating solutions to the problem statements presented above, a number of methods and results constituted novel contributions to their respective fields. The contributions were of two types: i) statistical methodology and ii) demonstration of new capability in X-ray imaging. These contributions are summarised in Table 1.1.

When developing the solutions to these problem statements, some methodology was developed that was found unnecessary or impractical for a software solution, which could be adopted by industry. As a major output of this work is practical solutions which are being adopted by industry, it is the actual practical solutions that are presented. Further methodology – which was unnecessary for satisfactory solutions to the problem statements but may be useful for analogous problems – are presented in the appendix.

The next chapter will introduce the inferential paradigm and the key apparatus used for the majority of modelling. In this thesis, I will use personal pronouns for my opinion and plural personal pronouns for the group involved in this research project.

Section	Contribution
2.3.1	A Bayes linear approach to making inferences about the segmentation of a mixture of multivariate random fields.
2.4	A method to approximate a Bayes linear adjustment with a kernel, enabling faster inference.
3.3	A closed form method for learning about the population mean of a set of second order exchangeable multivariate random fields.
3.4	Bayes linear methods for separately detecting high and low frequency anomalies.
3.4	Bayes linear methods for separately detecting spatial and spectral anomalies.
3.5.2	A method (termed continuous energy) for extracting multi-spectral material information from an X-ray system equipped with multi-absorption-plate (MAP) technology.
3.5.3	Demonstration of detecting plastic contaminants in chocolate using MAP technology.
4.2.3	A method for utilising signal processing methods in Bayesian emulation.
4.3.3	Demonstration of the inadequacy of non-stationary-isotropic PSF (current state-of-the-art) assumptions for scatter estimation.
4.3.4	Demonstration of the advantages of i) our non-stationary-non-isotropic method for scatter estimation over ii) current state-of-the-art.
5.2.1	A method for adjusting beliefs about random quantities with non-implausible sets.
5.2.1	A method for adjusting beliefs about multivariate random fields with non-implausible sets.
5.3	A method for history matching when beliefs about the inputs to a simulator are specified as multivariate random fields.
5.4	Iterative methods to find non-implausible points with fewer evaluations of the simulator.
5.6	Demonstration in simulation that Bayesian inverse problem solving may be used to make inferences about the material and thickness of an entity using mono-spectral X-ray examination technology.
5.6.1	Demonstration in simulation that iterative history matching can be used to find non-implausible points (the scatter estimates from which improve image quality) in a useful time-frame.
6.1.4	Extension of reified Bayesian analysis to the multiple systems case.
6.2	A method for tuning a simulator to its reified counterpart, given the direct tuning assumption.
6.3	Methods for i) assessing the direct tuning assumption and ii) making second order specifications for a) the random quantities involved in tuning and b) diagnostic quantities.
6.4	Demonstration on real world data that tuning can be used to make inferences about a reified simulator of GEANT4 that can be used to make direct inferences for reality.
6.6	Demonstration on real world data that the methods presented can be used to make inferences about the material and thickness of an entity from a mono-spectral X-ray image of that entity.
6.6	Demonstration on real world data that the methods presented can improve X-ray image quality.

Table 1.1: Notable novel contributions presented in this thesis.

Chapter 2

Bayes Linear Methods and Multivariate Random Fields

In this thesis, many of the quantities of interest may be organised as multivariate random fields. Therefore, a method for organising beliefs about multivariate random fields and adjusting those beliefs with vectors of elements on random fields is required. Therefore, the focus of this chapter is detailing the inferential paradigm that will be used in this thesis. Section 2.1 will discuss the inferential paradigm, Section 2.2 will present multivariate random fields and Section 2.2.2 will construct how inferences are made about multivariate random fields.

2.1 Bayes Linear Methods

A compelling description of a statistician's purpose is

> "to help people make better decisions under uncertainty" [105].

The subjective Bayesian paradigm is a flexible and coherent approach to achieving this purpose. In this section, I will develop the ideas of subjective Bayesian analysis, beginning with Bayes linear methods rather than the usual starting point of full Bayes. In the full Bayes approach, an expert's uncertainties for quantities of interest are encoded as probabilities and their beliefs about observables as conditional probabilities given the quantities of interest. Bayes' theorem provides the calculus for how their beliefs change when the values of uncertain quantities are revealed. Given that full Bayes is a special case of the Bayes linear approach, I take Bayes linear as primitive.

2.1.1 Subjective Bayesian Analysis

The Bayes linear approach is based on a logical set of ideas. Things in the world about which an expert is uncertain are assigned quantities. We elicit their uncertainties for these quantities and encode them as expectations and variances, and the relationship between the quantities as covariances. Some of the quantities may be observable and, when observed, they may change the expert's beliefs about the quantities of interest. When the values of observables are revealed, we find the linear combination of them which is as close as possible to viewing the quantities of interest themselves; this becomes the expert's adjusted belief.

In this description of subjective Bayesian inference, expectation rather than probability is taken as the primitive quantity. There are two major arguments in support of this. The first, presented by de Finetti in [94], is that when probability is primitive we must elicit an expert's probabilities for every possible event before any expectations may be calculated. When expectation is primitive, it is possible to only elicit expectations that are currently meaningful. Given that in many complex problems it is doubtful whether we can specify all possible outcomes, let alone elicit meaningful probabilities, it is sensible not to start from a position with unreasonable requirements. The second persuasive argument is that expectation offers a friendlier language in which beliefs may be discussed with non-statistical experts. Asking the question "What do you expect the outcome to be?" is more likely to be met with a well-thought-out response than "What is your probability that the outcome is x, for all x in the event space χ?". Given a statisticians purpose is to help people make better decisions under uncertainty, a clear advantage is a methodology which recognises their expertise without assuming statistical literacy. An informal definition of expectation is given by Jonathan Rougier in the lecture notes from his course on statistical inference [105].

> "My expectation of a random quantity X, denoted $\mathbb{E}[X]$, is my *best guess* for X, represented as a value in $\mathbb{R}$. In Statistics, unlike in probability theory, it is important to have some idea about what formal concepts actually mean, so that when I think about my 'expectation of sea-level rise in 2100' this conjures up a number in my mind. *Best guess* seems to work quite well."

This description is operational and provides a starting point to discuss uncertainties without the need for formal statistical training. The *best guess* description is, however,

ambiguous in that it is not clear what is meant by best. A more formal exposition is provided by de Finetti. If a random vector is notated $x = (x_{(1)}, \cdots, x_{(p)})^T$, the expectation of a random quantity $x_{(k)}$ is defined as the value an expert would choose for a if faced with the penalty

$$L = k(x_{(k)} - a)^2$$

when the value of $x_{(k)}$ is revealed. Here k is a constant defining units of loss. By this definition, the guess is best in the sense that it minimises a penalty. There may still be difficulties in eliciting meaningful expectations for all the components of a complex problem, but the physical meaning of expectation and its formal definition are clear.

Probability is then defined as the expectation of an indicator function for an event i.e.

$$\mathbb{P}[x_{(k)} = a] = \mathbb{E}[\mathbb{I}_{x(k)=a}]$$

where $\mathbb{I}_{x(k)=a} = 1$ if $x = a$ and 0 otherwise. Conditional expectation is defined as the value for a an expert would choose when faced with the penalty $k\mathbb{I}(x_{(k)} - a)^2$, where $\mathbb{I}$ is an indicator function for the conditioning event.

The variance of a random quantity, notated $\mathbb{V}ar[x]$, is defined in terms of expectations. It is an individual's expectation of the squared difference between a random quantity and their expectation for it.

Definition 2.1.1. ***Variance:*** *the variance of random vector x is*

$$\mathbb{V}ar[x] = \mathbb{E}\left[(x - \mathbb{E}[x])(x - \mathbb{E}[x])^T\right].$$

Variance is how far, in terms of squared difference, an expert thinks their *best guess* will be from the value when it is revealed. Variance informally specifies prior error bars or how far the revealed value would have to be from its expectation to cause surprise.

If y is a random vector $y = (y_{(1)}, \cdots, y_{(q)})^T$, the covariance between two uncertain quantities x, y is notated $\mathbb{C}ov[x, y]$. Covariance specifies beliefs about the linear relationship between two unknown quantities. It allows us to elicit beliefs such as; if x is larger than expected when revealed, then y is likely to be larger than expected as well.

Definition 2.1.2. ***Covariance:*** *the covariance between random vectors x and y is*

$$\mathbb{C}ov[x, y] = \mathbb{E}\left[(x - \mathbb{E}[x])(y - \mathbb{E}[y])^T\right].$$

Given a specification of beliefs, a calculus is required for how beliefs change when the values of uncertain quantities are revealed. In the Bayes linear approach, this is provided

by linear fitting and adjusted beliefs are referred to as adjusted expectations. The adjusted expectation for quantity $x_{(k)}$ with observation vector $z = (z_{(1)}, \cdots, z_{(l)})^T$ is defined as,

$$\mathbb{E}_z[x_{(k)}] = \beta_{(0)} + \sum_{i=1}^{l} \beta_{(i)} z_{(i)}$$

where $\beta = (\beta_{(0)}, \cdots, \beta_{(l)})^T$ minimises

$$\mathbb{E}\left[\left(x_{(k)} - \left(\beta_{(0)} + \sum_{i=1}^{p} \beta_{(i)} z_{(i)}\right)\right)^2\right].$$

This belief change calculus has an intuitive interpretation. We look for the linear combination of the data which is, in the sense of squared difference, expected to be as similar as possible to the quantity of interest; this is an individual's *best guess* for what that quantity is. The minimisation may be solved analytically, which leads to the following definition.

Definition 2.1.3. ***Adjusted expectation:*** *the adjusted expectation of x given z is*

$$\mathbb{E}_z[x] = \mathbb{E}[x] + \mathbb{C}ov[x, z]\mathbb{V}ar[z]^{-1}(z - \mathbb{E}[z]).$$

If $\mathbb{V}ar[z]$ is singular, the inverse is replaced with the Moore-Penrose generalised inverse. Adjusted variance and covariance are also key quantities when using Bayes linear methods for uncertainty quantification.

Definition 2.1.4. ***Adjusted Variance:*** *the adjusted variance of x given z is*

$$\mathbb{V}ar_z[x] = \mathbb{E}[(x - \mathbb{E}_z[x])(x - \mathbb{E}_z[x])^T] = \mathbb{V}ar[x] - \mathbb{C}ov[x, z]\mathbb{V}ar[z]^{-1}\mathbb{C}ov[z, x].$$

Definition 2.1.5. ***Adjusted Covariance:*** *the adjusted covariance of x and y given z is*

$$\mathbb{C}ov_z[x, y] = \mathbb{E}[(x - \mathbb{E}_z[x])(y - \mathbb{E}_z[y])^T] = \mathbb{C}ov[x, y] - \mathbb{C}ov[x, z]\mathbb{V}ar[z]^{-1}\mathbb{C}ov[z, y].$$

Therefore, to perform a Bayes Linear analysis with random vectors $\{x, y, z\}$, a second order specification is required.

Definition 2.1.6. ***Second order specification:*** *the second order specification of $\{x, y, z\}$ is*

$$\mathbb{E}[(x^T, y^T, z^T)^T], \mathbb{V}ar[(x^T, y^T, z^T)^T].$$

A special case of adjusted expectation is when x is an indicator function and z forms a partition i.e. $z_{(i)}$ is either zero or one and $\sum_i z_{(i)} = 1$. Then

$$\mathbb{E}_z[\mathbb{I}_{x(k)=a}] = \mathbb{P}(x_{(k)}|z) = \sum_i \mathbb{P}(x_{(k)}|z_{(i)}) z_{(i)}$$

which − if performed for all possible values of $x_{(k)}$ − is equivalent to the full Bayesian analysis.

Bayes linear is a coherent formulation of subjective Bayesian inference when expectation is primitive and it has been shown to be useful in many applications (see, for example, [22, 27, 32, 117]). Taking i) linear fitting as the primitive calculus for belief change and ii) conditioning derived as a special case is complementary to taking i) expectation as the primitive quantity for belief specification and ii) probability derived as a special case. It has advantages over limiting analyses to the full Bayes approach: the computations are often easier and only a second order specification is required before inference may be performed. For a detailed exposition of the Bayes linear approach, see [28].

Properties of Adjusted Expectations

Adjusted expectations have a number of properties which are useful when quantifying uncertainty. Adjusted expectations are linear, meaning that for any two p dimensional random vectors x, x' and conformable matrices a, b

$$\mathbb{E}_z[ax + by] = a\mathbb{E}_z[x] + b\mathbb{E}_z[x'].$$

Adjusted expectations are also conglomerable, meaning

$$\mathbb{E}[\mathbb{E}_z[x]] = \mathbb{E}[x].$$

The difference between a quantity and its adjusted expectation also has useful properties and plays a pivotal role in uncertainty quantification. This is defined as the adjusted version.

Definition 2.1.7. ***Adjusted version:*** *the adjusted version of x given z is*

$$\mathbb{A}_z[x] = x - \mathbb{E}_z[x].$$

Adjusted versions have the following second order properties:

$$\mathbb{E}[\mathbb{A}_z[x]] = 0,\ \mathbb{V}ar[\mathbb{A}_z[x]] = \mathbb{V}ar_z[x],\ \mathbb{C}ov[\mathbb{A}_z[x], z] = 0 \text{ and } \mathbb{C}ov[\mathbb{A}_z[x], \mathbb{E}_z[x]] = 0.$$

Given definition 2.1.7, x may be rewritten as

$$x = \mathbb{A}_z[x] + \mathbb{E}_z[x].$$

Definition 2.1.8. ***Resolved variance:*** *the variance of x resolved by z is*

$$\mathbb{RV}ar_z[x] = \mathbb{V}ar[\mathbb{E}_z[x]] = \mathbb{C}ov[x, z]\mathbb{V}ar[z]^{-1}\mathbb{C}ov[z, x].$$

Given definition 2.1.8, the variance of x may be decomposed as

$$\mathbb{V}ar[x] = \mathbb{V}ar_z[x] + \mathbb{V}ar[\mathbb{E}_z[x]] = \mathbb{V}ar_z[x] + \mathbb{R}\mathbb{V}ar_z[x].$$

In other words, the prior variance of a random quantity is equal to the variance of the adjusted version plus the variance resolved by the adjustment. This decomposition in comfortingly intuitive; the uncertainty in a random quantity is equal to the uncertainty remaining after adjustment plus the uncertainty resolved by the adjustment.

Practising Bayes Linear

The practical application of Bayes linear methods is separated into three phases:

1. elicit a second order specification for all random quantities involved and analyse the expected adjustments,
2. interpret the actual adjustments and
3. analyse the difference between actual and expected adjustments.

In this thesis, three statistics are used for the majority of belief analysis: discrepancy, standardised errors and transformed standardised errors.

1. The first quantity is is based on the Mahalanobis distance, see [106].

   Definition 2.1.9. ***Discrepancy:*** *the discrepancy of a random vector x is*

   $$Dis(x) = (x - \mathbb{E}[x])^T(\mathbb{V}ar[x])^{-1}(x - \mathbb{E}[x]).$$

   Given an appropriate belief specification, discrepancy has prior expectation equal to the rank of the matrix $\mathbb{V}ar[x]$. The prior variance specification depends on the distribution of x. For example, if an expert asserts the distribution of x is multivariate Gaussian, then their variance of the discrepancy measure is given by $2\text{rank}(\mathbb{V}ar[x])$.

2. The second quantity is similar to standardised residuals in traditional statistical inference.

   Definition 2.1.10. ***Standardised errors:*** *the standardised error of a random quantity $x_{(k)}$ is*

   $$Std(x_{(k)}) = \frac{x_{(k)} - \mathbb{E}[x_{(k)}]}{\sqrt{\mathbb{V}ar[x_{(k)}]}}. \tag{2.1.1}$$

Standardised errors have prior expectation zero and prior variance one.

3. The third diagnostic assesses the covariance specification of a quantity.

Definition 2.1.11. ***Transformed standardised errors.*** *Given* $LL^T = \mathbb{V}ar[x]$, *the transformed standardised errors of a random vector* x *is*

$$Std_T(x) = L^{-1}x$$

where L *is unique if* $\mathbb{V}ar[x]$ *is positive-definite [119].*

Transformed standardised errors have prior expectation zero and prior variance equal to an identity matrix I_p the same size as $\mathbb{V}ar[x]$.

These statistics measure the difference between expected and revealed, scaled by uncertainty. Given an appropriate belief specification, their second order specification is known, which allows an assessment of whether the revealed value is significantly larger or smaller than expected. To decide if the value is *large* or *not large*, probability heuristics are employed. For example, as $\mathbb{P}[|Std[x_{(k)}]| > 3] < 0.05$ by the three sigma rule [47], standardised residuals greater than three or less than minus three could indicate further consideration of the specification is required. Diagnostic measures give a systematic way of assessing if the specification can be taken at face value and draw attention to the components of the specification which require attention. A failure of diagnostics is not in itself a reason to throw out an adjusted belief; it indicates that we should think carefully about whether the specifications are in fact what an individual believes and whether the calculus of belief adjustment reflects their current position.

2.2 Multivariate Random Fields and Observations

It is assumed that y is a vector of elements on a multivariate random field, meaning that the elements of y may be grouped according to a meaningful indexing of its underlying domain. For example, y could consist of multiple readings at multiple spatial locations. Then all observations at a given location are grouped together and each group is given a spatial index.

Definition 2.2.1. ***Multivariate random field.*** *Given a normed-vector-space* $\mathcal{T} \subset \mathbb{R}^p$, *a multivariate random field is defined as a set of* q *dimensional random variables*

$$\{y(t) \ : \ t \in \mathcal{T}\}$$

which may be indexed by elements $t \in \mathcal{T}$.

Random fields are useful as specifying beliefs in terms of t can be easier for complex constructs. A set of random fields is notated

$$\{y^{(1)}, \cdots, y^{(N)}\}.$$

So the superscript notates the multivariate random field within a set. The set of locations on a field in y is notated $T = (t^{(1)}, \cdots, t^{(n)})$ where $t = (t_{(1)}, \cdots, t_{(p)})^T$. The number of observations at each location is q. Given the definition of a multivariate random field, y is organised as

$$y = (y(t^{(1)}), \cdots, y(t^{(n)}))^T \text{ where} \tag{2.2.1}$$

$$y(t) = (y_{(1)}(t), \cdots, y_{(q)}(t))^T. \tag{2.2.2}$$

Therefore, an element from a set of multivariate random fields is notated

$$y^{(i)}_{(j)}(t) = y^{(\text{multivariate random field } i)}_{(\text{constituent field } j)}(\text{location } t).$$

A summary of the notation used in this thesis is given in Appendix A.1.

Observations

S is defined as the entity about which an expert would like to make inferences. For example, S could be a chocolate bar, a patient, a piece of software or a securities market for a particular time period. The process of making an error-free observation of S is given by

$$y = f(S, x^f_E) \tag{2.2.3}$$

where f is a function describing the experimental equipment and x^f_E is a vector describing experimental decisions, post-processing parameters and conditions when the observation is made. An error-free observation is the measurement that would be made if there was complete control over the experimental set up and enough resources to meaningfully resolve any stochastic error with averaging. As such, it is viewed as a deterministic functional relationship. In x^f_E, the superscript is to indicate its connection with examination process f and subscript to denote association with experimental parameters. This will be useful in Chapter 6 when linking inputs to a simulator and x^f_E.

S may be continuous and y is a vector values at discrete locations on a random field. This connection may be as simple as y being a set of locations on S but often it will be more

complex. The relationship between the two is important to making meaningful inferences about S. However, it is less important in the anomaly detection problem (see Chapter 3) as the primary interest is in the population to which S belongs not the composition of S itself. Further discussion on this point will therefore be reserved for Chapter 4 after which it plays a larger role.

The sample set $\mathcal{S} = \{S^{(1)}, \cdots, S^{(N)}\}$ is measured and the observation training-set is notated $D^f = \left\{y^{(1)}, \cdots, y^{(N)}\right\}$ where $y^{(i)} = f(S^{(i)}, x_E^{f(i)})$. In this thesis, it will be assumed that the error-free experimental set up is constant (i.e. $x_E^{f(i)} = x_E^{f(j)} \; \forall \; i, j$), but, depending on the application, it may be necessary to make it variable.

2.2.1 Decomposing Sources of Uncertainty

When specifying beliefs about random fields in complex problems, it is useful to decompose the complete field into a linear combination of random fields describing separate sources of uncertainty. This allows the complete specification to be made in terms of the random fields describing the constituent sources of uncertainty. This approach is common in probabilistic modelling of complex problems; see, for example, Bayesian networks in [97]. For example, in the plastic in chocolate problem (see problem statement 1.1.1), three sources of uncertainty were considered: measurement error, uncertainty in the population mean and uncertainty in the population variance.

Measurement Error

Measurement error encodes the belief that if the same entity is measured multiple times, the result will not necessarily be identical. Here the measurement error is assumed to take the form of additive noise uncorrelated with y i.e.

$$z = y \oplus \epsilon_{me}$$

where ϵ_{me} is assumed to be an expectation zero random field describing measurement error. The notation $\oplus$ is to indicate the sum of two *a-priori* uncorrelated random quantities. Measurement error may occur because the examination process is inherently random or because the experimental conditions are difficult to stabilise. Consideration of these two sources of measurement error implies that the decomposition of ϵ_{me} into two uncorrelated components, describing local and global variation, will aid specification. The local part encodes the belief that knowledge of measurement error at a location on the random field

will only change beliefs about the measurement error in a local neighbourhood. The global component encodes the belief that knowledge of the error at a location on a random field will change beliefs about the error everywhere on that field. It is natural to expect that changes in experimental conditions will effect the entire multivariate field and, as such, would constitute global variation. ϵ_{me} is rewritten as

$$\epsilon_{me} = \epsilon_{me}^{G} \oplus \epsilon_{me}^{L}$$

where ϵ_{me}^{G} describes the global variation and ϵ_{me}^{L} the local variation.

Population Mean

The second source of uncertainty is about the population mean. It is assumed that each entity is a sample from an underlying population and that each sample may be written as the uncorrelated sum of the population mean and a residual from that mean i.e.

$$S^{(i)} = M(S) \oplus R^{(i)}(S).$$

Here $M(S)$ is the population mean and $R^{(i)}(S)$ is the residual from that mean. $R^{(i)}(S)$ is assumed to have expectation zero, $\mathbb{V}ar[R^{(i)}(S)] = \mathbb{V}ar[R^{(j)}(S)]$ and $\mathbb{C}ov[R^{(i)}(S), R^{(j)}(S)] = 0$. This construction will also be formulated as a result of the assumption that the samples form part of a second order exchangeable sequence of random variables, however, discussion on this point is reserved until Section 3.2. Given that the entities have underlying variation, it is logical to expect that this will induce variation in the error-free observations. It is assumed that they may be decomposed as

$$f(S^{(i)}, x_E^f) = y^{(i)} = M(y) \oplus R^{(i)}(y)$$

where $\mathbb{E}[R^{(i)}(y)] = 0$, $\mathbb{V}ar[R^{(i)}(y)] = \mathbb{V}ar[R^{(j)}(y)]$ and $\mathbb{C}ov[R^{(i)}(y), R^{(j)}(y)] = 0$. The population mean may be a complex field about which it is practically difficult to specify beliefs. A tractable approach to encoding this complexity is to decompose the specification into global and local components (as used in response surface methodology [91, 92]). The global component captures the major trend while the local component deals with higher frequency behaviour. When this decomposition is appropriate, the population mean is rewritten as

$$M(y) = M^{G}(y) \oplus M^{L}(y)$$

where $M^{L}(y)$ is a random field describing local variation and $M^{G}(y)$ describes global variation.

Population Variance

The third source of uncertainty is in the population variance. Elements of $R^{(i)}(y)$ may be correlated. For example, if the sample of interest is a confectionery bar made of chocolate and wafer, then the wafer will imprint an oscillatory component on the observation. In this case, knowledge of the residual from the population mean at one element would change beliefs about another. These beliefs are included by decomposing the residual into uncorrelated components describing local and global variation. The residual from the population mean is modelled as

$$R^{(i)}(y) = R^{G(i)}(y) \oplus R^{L(i)}(y).$$

To summarise the model for a multivariate random field in anomaly detection, an expert may be uncertain about the quantity of interest because they are uncertain about the population mean, the residual from the population mean and the measurement error on an observation. This is modelled by the independent sum of six random variables: three of which describe global variation and three of which describe local variation. The model for a multivariate random field encoding beliefs about an observation for the anomaly detection problem is

$$z^{(i)} = M^G(y) \oplus M^L(y) \oplus R^{G(i)}(y) \oplus R^{L(i)}(y) \oplus \epsilon_{me}^{G(i)} \oplus \epsilon_{me}^{L(i)}. \tag{2.2.4}$$

2.2.2 Inference about Multivariate Random Fields

The random variables encoding uncertainties about physical quantities of interest are specified as multivariate random fields. The observations are assumed to be a discrete set of locations on a multivariate random field. Therefore, a methodology is required for specifying beliefs about random fields and adjusting those beliefs with a set of locations on a field. In this section, a *toolbox* of techniques which are tractable and flexible will be presented. The methods share a similar formalism to Bayesian emulation, which has already been shown to be effective in modelling complex computer models, thus enabling the wealth of methodologies available for emulation to be applied to multivariate random fields.

General Form

In this work, beliefs about an element on a multivariate random fields is organised as a sum of two *a-priori* uncorrelated random fields

$$y_{(k)}(t) = \mathcal{M}_{(k)}(t) \oplus \mathcal{W}_{(k)}(t). \tag{2.2.5}$$

Here $\mathcal{M}_{(k)}(t)$ is referred to as the mean function which describes global variation while $\mathcal{W}_{(k)}(t)$ is known as the residual process which describes local variation. It is also assumed that the residual process has prior expectation zero. A particularly tractable form of the specification is when the mean function is taken to be a linear combination of basis functions with associated regression parameters. Furthermore, the same form of the mean function is assumed to be sufficient for all fields. This form is given by

$$y_{(k)}(t) = \sum_{m=1}^{r} \beta_{(m,k)} h_{(m)}(t) \oplus \mathcal{W}_{(k)}(t) = h(t)\beta_{(:,k)} \oplus \mathcal{W}_{(k)}(t) \tag{2.2.6}$$

where $h(t)$ is a $(1 \times r)$ vector of known basis functions and β is a $(r \times q)$ matrix of random quantities referred to as regression parameters. To perform a Bayes linear adjustment, a second order specification is required. The second order specification for the mean function is equivalent to a second order specification for β. $\mathcal{W}(t)$ is assumed to be a weakly stationary multivariate random field with second order specification made using a covariance function.

Covariance Functions

A covariance function is an artefact to organise the specification of covariance between any two locations on a random field. The general form is given by

$$\mathbb{C}ov[\mathcal{W}_{(k)}(t), \mathcal{W}_{(k')}(t')] = \sqrt{\sigma_{(k)}(t)\sigma_{(k')}(t')}C(t, t', k, k').$$

Here $C(t, t', k, k')$ is a correlation function and $\sigma_{(k)}(t)$ is a variance function. This splits up the covariance specification into variance and correlation specifications, which is usually easier. A weakly stationary covariance function is when

$$\sigma_{(k)}(t) = \sigma_{(k)} \; \forall \, t \; \in \mathcal{T} \text{ and}$$
$$C(t, t', k, k') = C(t + \epsilon, t' + \epsilon, k, k') \; \forall \epsilon, t, t' \; \in \mathbb{R}^p$$

An isotropic correlation function is when the covariance specification is invariant to rotations i.e. of the form

$$C(t, t', k, k') = C(Ut, Ut', k, k') \; \forall t, t' \in \mathbb{R}^p \text{ and } U \in SO(p).$$

Here U is a $p \times p$ matrix and $SO(p)$ is the special-orthogonal-group of order p [125]. Therefore, a weakly stationary and isotropic covariance function is when

$$C(t, t', k, k') = C_0(|t - t'|, k, k') \; \forall \; t, t' \; \in \mathbb{R}^p$$

holds for some function C_0. A tractable approach to specifying $C(t, t', k, k')$ is to assume separability.

Definition 2.2.2. ***Separable covariance function:*** *A separable covariance specification is given by*

$$C(t, t', k, k') = c_{(k,k')} C(t, t'; l) \tag{2.2.7}$$

where $c_{(k,k')} = 1$ *if* $k = k'$.

l is the correlation length which describes how this specification changes with distance and $c_{(k,k')}$ is the correlation between fields. In this thesis, each specification has a fixed correlation length, so $C(t, t'; l)$ notates that t, t' vary while l is fixed. The correlation length can be interpreted as controlling the smoothness of the residual process; for most correlation functions, the larger the correlation length the smoother the field. Given separability, the second order specification of the residual process is given by

$$\mathbb{C}ov[\mathcal{W}(t), \mathcal{W}(t')] = \Sigma \otimes C(t, t'; l)$$

where $\otimes$ denotes a Kronecker product (see [29]) and $\Sigma_{(k,k')} = \sqrt{\sigma_{(k)}\sigma_{(k')}} c_{(k,k')}$.

Definition 2.2.3. ***Kronecker product:*** *the Kronecker product of two matrices A and B is given by*

$$A \otimes B = \begin{bmatrix} A_{(1,1)}B & A_{(1,2)}B & \cdots & A_{(1,m)}B \\ \vdots & \vdots & \ldots & \vdots \\ A_{(n,1)}B & A_{(n,1)}B & \cdots & A_{(n,m)}B \end{bmatrix}$$

This construction has computational advantages, the specification is relatively simple and the process of adjusting the hyper-parameter Σ is tractable (see key-formula 2.2.1). The downside is that there is an implicit assumption that the residual process for every field is equally smooth. The second approach used in this thesis is to model each of the q fields separately with a distinct correlation length, given by

$$C(t, t', k, k) = C(t, t'; l_{(k)}) \text{ and}$$
$$C(t, t', k, k') = 0 \text{ for } k \neq k'.$$

In our experience, the correlation in t has a larger impact on the inference than the correlation between fields. Therefore, treating the fields separately is the better starting place. The form of correlation function used in this work is the exponential family.

Definition 2.2.4. ***Exponential family covariance functions:*** *the exponential family of covariance functions is*

$$C(t, t'; l, p) = \exp\left[-\sum_j \frac{(|t_{(j)} - t_{(j')}|)^p}{2l_{(j)}^p}\right]. \tag{2.2.8}$$

The notation $C(t, t'; l, p)$ is to notate that for a given specification, l and p are fixed while t, t' vary. $p \in [1, 2]$ works well with $p = 2$, referred to as the squared exponential, reflecting the belief that the residual process is smooth (infinitely differentiable) and $p < 2$ that the residual process is rough (non-differentiable). The left hand plot in Figure 2.1 shows the effect of p on the covariance functions. The right hand plot shows the effect of changing p on samples from a Gaussian process.

Definition 2.2.5. ***Gaussian Process:*** *a random field where all finite sets of locations on the field are multivariate Gaussian.*

This plot illustrates how p may be changed to correspond to prior beliefs about the smoothness of a field. For further discussion on correlation functions, see [30] Chapter 4 or [44].

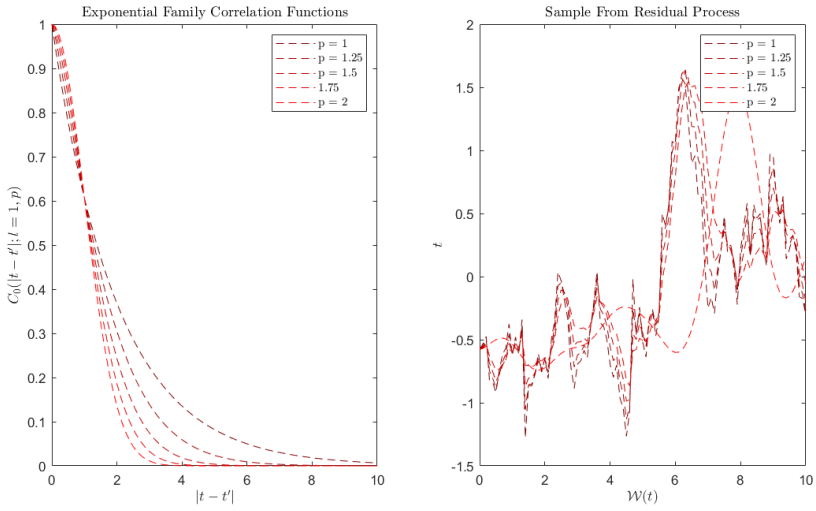

Figure 2.1: Left: exponential family covariance functions for varying p. Right: samples from a zero mean and variance one Gaussian process for varying p.

Adjusting Beliefs About Multivariate Random Fields

For the form of multivariate random fields given in equation 2.2.6, the required second order specification necessary for a Bayes linear adjustment of y with z is

$$\mathbb{E}[\beta_{(:)}] = \mu_\beta \ , \ \mathbb{V}ar[\beta_{(:)}] = V_\beta \ , \ \mathbb{V}ar[\mathcal{W}_{(:)}(T)] = V_\mathcal{W} \text{ and } \mathbb{V}ar[\epsilon_{me(:)}] = V_{me}.$$

The second order specification for the observation z is

$$\begin{aligned}\mathbb{E}[y_{(:)}] = \tilde{H}\mu_\beta = \mathbb{E}[z_{(:)}] = \mu_z \text{ and}\\ \mathbb{V}ar[z_{(:)}] = \tilde{H}V_\beta\tilde{H}^T + V_\mathcal{W} + V_{me} = V_z.\end{aligned}$$

Here $\tilde{H} = (I_q \otimes H)$ and $H = h(T) = (h(t^{(1)})^T, \cdots, h(t^{(n)})^T)^T$ is the design matrix. The adjusted beliefs at location t' are given by

$$\begin{aligned}\mathbb{E}_z[y(t')] &= \tilde{h}(t')\mathbb{E}_z[\beta_{(:)}] + \mathbb{E}_z[\mathcal{W}(t')] \ ,\\ \mathbb{V}ar_z[y(t')] &= \tilde{h}(t')\mathbb{V}ar_z[\beta_{(:)}]\tilde{h}(t')^T + \mathbb{V}ar_z[\mathcal{W}(t')] + \tilde{h}(t')\mathbb{C}ov_z[\beta_{(:)}, \mathcal{W}(t')]\\ &\quad + \mathbb{C}ov_z[\mathcal{W}(t'), \beta_{(:)}]\tilde{h}(t')^T \text{ and}\\ \mathbb{C}ov_z[y(t'), y(t^*)] &= \tilde{h}(t')\mathbb{V}ar_z[\beta_{(:)}]\tilde{h}(t^*)^T + \mathbb{C}ov_z[\mathcal{W}(t'), \mathcal{W}(t^*)]\\ &\quad + \tilde{h}(t')\mathbb{C}ov_z[\beta_{(:)}, \mathcal{W}(t^*)] + \mathbb{C}ov_z[\mathcal{W}(t'), \beta_{(:)}]\tilde{h}(t^*)^T\end{aligned}$$

where $\tilde{h}(t) = (I_q \otimes h(t))$. Then given $\mathbb{C}(a, b) = \mathbb{C}ov[\mathcal{W}_{(:)}(a), \mathcal{W}_{(:)}(b)]$,

$$\begin{aligned}\mathbb{E}_z[\beta_{(:)}] &= \mu_\beta + V_\beta\tilde{H}^T V_z^{-1}(z_{(:)} - \mu_z),\\ \mathbb{V}ar_z[\beta_{(:)}] &= V_\beta - V_\beta\tilde{H}^T V_z^{-1}\tilde{H}V_\beta,\\ \mathbb{E}_z[\mathcal{W}_{(:)}(t')] &= \mathbb{C}(t, T)V_z^{-1}(z_{(:)} - \mu_z) \ ,\\ \mathbb{V}ar_z[\mathcal{W}_{(:)}(t')] &= \mathbb{C}(t', t') - \mathbb{C}(t', T)V_z^{-1}\mathbb{C}(T, t'),\\ \mathbb{C}ov_z[\mathcal{W}_{(:)}(t), \beta_{(:)}] &= -\mathbb{C}(t, T)V_z^{-1}\tilde{H}V_\beta \text{ and}\\ \mathbb{C}ov_z[\mathcal{W}_{(:)}(t'), \mathcal{W}_{(:)}(t^*)] &= \mathbb{C}(t', t^*) - \mathbb{C}(t', T)V_z^{-1}\mathbb{C}(T, t^*).\end{aligned}$$

Diagnostics for Multivariate Random Fields

Given beliefs have been specified and adjusted, it is important to assess whether any of the specifications require further consideration. This process is called validation and it is carried out using diagnostic measures which are designed such that, given an appropriate update of an expert's beliefs, the second order specification is known. In other words, if a random field is sampled from a distribution with the asserted second order specification

and adjusted beliefs calculated, the diagnostics have known second order properties. Deviation from this specification is evidence to suggest that elicitation may require further consideration before the adjusted beliefs may be taken at face value.

The form presented for modelling multivariate random fields is flexible and can encode a wide variety of prior beliefs. However, poor predictive performance can be caused by inappropriate specification of the mean function, residual process correlation structure or hyper-parameters. The mean function may either have too many, too few or poorly chosen basis functions. Too few or poorly chosen basis functions will result in no account being taken of a proportion of the non-stationary behaviour and the stationary-isotropic residual process assumptions will then be inappropriate. Too many will result in too much local variation being resolved by the mean function. The residual process assumptions may also be unjustified if the variance structure is heteroscedastic. Finally, even if the specifications are appropriate, there are a number of hyper-parameters which have to be specified or adjusted. These parameters are unfamiliar to non-statisticians and, therefore, it may be unreasonable to expect a meaningful belief specification.

An inappropriate stationary-isotropic residual process assumption will cause uncertainty to be underestimated at some locations and overestimated at others. Too many basis functions will lead to overconfidence everywhere and inappropriate hyper-parameters can lead to either underconfidence or overconfidence at all locations on the field.

In order to validate predictive performance, a data-set not included in the inference is useful. A portion of the data-set can be used for validation. The validation set is never included in the inference and the diagnostic quantities are calculated for those observations. When there is no validation set available, cross-validation can be used instead. Cross validation selects a group of elements from the training-set and uses them as a validation set. The process is then repeated with a new distinct group until every element has been removed at least once. It is assumed that the observation vector is split into two, $\{z(T'), z(T^*)\}$. Then three diagnostic quantities are used with a scatter plots of t against the diagnostic quantities. Histograms of the diagnostic quantities are also useful for diagnosing the nature of a problem.

1. The first diagnostic is the standardised prediction error, given by

$$Std\left(z_{(k)}(t') - \mathbb{E}_{z(T^*)}[z_{(k)}(t')]\right) = \frac{z_{(k)}(t') - \mathbb{E}_{z(T^*)}[z_{(k)}(t')]}{\sqrt{\mathbb{V}ar_{z(T^*)}[z_{(k)}(t')]}}$$

where $t' \in T'$. This quantity has prior expectation zero and variance one. Pukelsheim's

three sigma rule [47] can be used to diagnose a problem. For example, standardised residuals greater than three indicate a possible conflict between adjusted beliefs and validation observations. As the variance of the transformed residual should be homoscedastic, large numbers of very small values complemented with a few between one and three would also imply a problem.

The scatter plots of standardised residuals against t should look like correlated samples from a distribution with expectation zero and variance one. The histogram should be uni-modal and symmetric about zero, with the majority of the mass between minus three and minus three. Uni-modality is not a necessary requirement of the Bayes linear approach. However, without these features adjusted beliefs have questionable value.

2. The second quantity is prediction discrepancy, given by

$$Dis\left(z(t') - \mathbb{E}_{z(T^*)}[z(t')]\right) = \left(z(t') - \mathbb{E}_{z(T^*)}[z(t')]\right)^T \left(\mathbb{V}ar_{z(T^*)}[z(t')]]\right)^{-1} \left(z(t') - \mathbb{E}_{z(T^*)}[z(t')]\right)$$

with prior expectation q and variance dependent on distributional assumptions. Very large values or too many small values indicate cause for concern. However, given the variance is dependent on possibly unspecified distributions, spotting this is more difficult. A good starting rule is values over $q + 3\sqrt{2q}$ indicates a possible problem. This is essentially making Gaussian assumptions for all quantities involved, which implies $\mathbb{V}ar\left[Dis\left(z(t') - \mathbb{E}_{z(T^*)}[z(t')]\right)\right] = 2q$.

Given Gaussian assumptions, the histogram of the discrepancy should look unimodal and approximately chi-squared with q degrees of freedom. The scatter plots of discrepancy can show local but not global correlation.

3. The third quantity is transformed standardised errors which diagnose whether the posterior correlation structure is appropriate. If $\mathbb{V}ar_{z(T^*)}[z_{(:)}(T')] = LL^T$, then the quantity is

$$Std_T\left(z_{(:)}(T') - \mathbb{E}_{z(T^*)}[z_{(:)}(T')]\right) = L^{-1}\left(z_{(:)}(T') - \mathbb{E}_{z(T^*)}[z_{(:)}(T')]\right) \tag{2.2.9}$$

where L may be calculated using, for example, the Cholesky decomposition. Again, by Pukelsheim, values over three indicates cause for concern but most importantly the transformed prediction errors should be uncorrelated.

The scatter plot of transformed residuals should look like independent samples from a distribution with expectation zero and variance one. The histogram should be uni-modal and symmetric about zero.

When the stationary-isotropic residual process assumption is inappropriate, there may be i) uneven variance in the scatter plots, ii) outliers or correlation in the transformed residuals or iii) multi-modal histograms. When the mean function has too many basis functions there may be i) clear waves in the scatter plots, ii) correlation in the transformed standardised errors and iii) overly wide histograms. The correlation length being too small will cause the histograms to be too narrow while inappropriate variance hyper-parameters will exhibit histograms which are either too wide or too narrow.

Cross validation is in our experience the most effective validation procedure. However, it is expensive as validation must be repeated many times. For more details on cross validation, see [93] and [58].

2.2.3 Generalised Least Squares Point Estimates (GLS)

If the second order specification is too challenging or data is plentiful enough to negate the prior, the generalised least square estimates may be used to approximate adjusted beliefs. This is analogous to a weak prior for Bayes linear methods and in practice it is often an effective starting point. Suppose the aim is to calculate adjusted beliefs about y by an observation vector $z = y + \epsilon_{me}$. Given the separable residual covariance structure in equation 2.2.7, $D = \text{Corr}[\mathcal{W}_{(k)}(T)] + \text{diag}(w(T))I_n$ is defined as the stationary variance specification, where $\text{diag}(x)$ is a function which returns a square matrix with diagonal elements x. $w(t) = \frac{\sigma_{me(k,k)}(t)}{\Sigma_{(k,k)}}$ is the ratio of uncertainty in the residual process and uncertainty induced by measurement error and it is assumed to be constant for all k. The GLS estimates for the regression parameters and variance multipliers are then defined as

$$\beta^* = (H^T D^{-1} H)^{-1} H^T D^{-1} z \text{ and} \tag{2.2.10}$$

$$\Sigma^* = \frac{1}{n-r}(z - H\beta^*)^T D^{-1}(z - H\beta^*). \tag{2.2.11}$$

For multivariate random fields, this is assuming that w is the same for all fields. Then the adjusted beliefs are approximated by

$$\begin{aligned}\mathbb{E}_z[y(t')] &\approx h(t')\beta^* + \mathbb{C}(t',T)D^{-1}(z - H\beta^*) = h(t')\beta^* + \mathcal{W}^*(t')\\ \mathbb{V}ar_z[y(t')] &\approx \Sigma^* \otimes (1 - \mathbb{C}(t',T)D^{-1}\mathbb{C}(T,t')\\ &\quad + (h(t') - \mathbb{C}(t',T)D^{-1}H)(H^TD^{-1}H)^{-1}(h(t') - \mathbb{C}(t',T)D^{-1}H)^T)\end{aligned}$$

where $\mathbb{C}(a,b) = \text{Corr}(\mathcal{W}_{(k)}(a), \mathcal{W}_{(k)}(b))$. This approach is analogous to assuming the field is a Gaussian process and weak prior assumptions for β and Σ, see [30, 59] for details. It has large computational advantages as it scales well as the number of constituent fields increases. The two assumptions are that the correlation length and w are appropriate for all fields, which may not be the case and, if it is not, separate GLS models for each field work well.

Key-formula 2.2.1. ***GLS estimates:*** *the GLS estimates for adjusted beliefs of y given z are*

$$\begin{aligned}D &= Corr[\mathcal{W}_{(k)}(T)] + diag(w(T))I_n,\\ \beta^* &= (H^TD^{-1}H)^{-1}D^{-1}z,\\ \Sigma^* &= \frac{1}{n-r}(z - H\beta^*)^TD^{-1}(z - H\beta^*),\\ \mathbb{E}_z[\mathcal{W}(t')] \approx \mathcal{W}^*(t') &= \mathbb{C}(t',T)D^{-1}(z - H\beta^*),\\ \mathbb{E}_z[\beta] &\approx \beta^*,\\ \mathbb{V}ar_z[\mathcal{W}(t')] \approx \mathbb{V}ar[\mathcal{W}^*(t') - \mathcal{W}(t')] &= \Sigma^* \otimes (1 - \mathbb{C}(t',T)D^{-1}\mathbb{C}(T,t')),\\ \mathbb{V}ar_z[\beta_{(:)}] \approx \mathbb{V}ar[\beta^*_{(:)} - \beta_{(:)}] &= \Sigma^* \otimes \left(H^TD^{-1}H\right)^{-1},\\ \mathbb{E}_z[y(t')] &\approx h(t')\beta^* + \mathcal{W}^*(t'),\\ \mathbb{V}ar_z[y(t')] &\approx \Sigma^* \otimes (1 - \mathbb{C}(t',T)D^{-1}\mathbb{C}(T,t')\\ &\quad + (h(t') - \mathbb{C}(t',T)D^{-1}H)(H^TD^{-1}H)^{-1}\\ &\quad (h(t') - \mathbb{C}(t',T)D^{-1}H)^T).\end{aligned}$$

2.3 Mixtures of Multivariate Random Fields

The form of a multivariate random field, given in equation 2.2.6, performs well when the global variation is differentiable (referred to as mean-differentiable). However, for mean-non-differentiable fields, it can perform badly, falling somewhere between two extremes:

1. the smooth (mean-differentiable) regions are too rough, leading to an overestimation of uncertainty and

2. the edge (mean-non-differentiable) regions are too smooth, leading to an underestimation of uncertainty.

Given that many applications will involve a mean-non-differentiable field an extension – with more flexibility while retaining most of the tractability – is pertinent. The general form for a multivariate random field is modified to a mixture of multivariate random fields where each component of the mixture is specified as a mean-differentiable random field with stationary-isotropic residual correlation structure.

The space is partitioned such that beliefs about elements within a partition may be specified as a random field of the form given by equation 2.2.6. Each partition will be notated $\mathcal{T}_\omega$ and the set of partitions have the properties

$$\cup_{\omega=1}^{\tilde{q}} \mathcal{T}_\omega = \mathcal{T} \text{ and } \mathcal{T}_\omega \cap \mathcal{T}_{\omega'} = \emptyset \; \forall \omega \neq \omega'.$$

The form of the multivariate random field becomes

$$y_{(k)}(t) = \sum_{\omega=1}^{\tilde{q}} \pi_{(\omega)}(t) y_{(k,\omega)}(t). \tag{2.3.1}$$

The subscripts (k, ω) notate (*constituent field*, *constituent partition*). $\pi_{(\omega)}(t)$ is an indicator field meaning $\pi_{(\omega)}(t) = 1$ if $t \in \mathcal{T}_\omega$ and 0 otherwise. $y_{(k,\omega)}(t)$ is the random field for elements belonging to partition ω. In this thesis, the number of partitions is assumed to be known. This type of model is useful in many imaging applications as images often exhibit smooth areas broken up by discontinuities; referred to as a plateau and ridge structure.

A large amount of additional flexibility is added by the mixtures model. In this thesis, it is assumed that the number of partitions is known and the location of the control point does not change the specification of the constituent of the mixture. This is because the examples considered have a well known shape and the control points are within a *small* neighbourhood of their expected location. For example, consider an X-ray image of a chocolate bar with a wafer inside. One partition models the region with wafer while the other models the partition with pure chocolate. The wafer may move a small amount within the chocolate and how the wafer has moved does not affect how it is modelled. Therefore, although a simple case, it is appropriate for the industrial examples presented.

The partitions can be specified directly or using control points and a partitioning function.

Definition 2.3.1. ***Control point:*** *a control point $\rho^{(i)}$ is a location in $\mathcal{T}$ and any location in $\mathcal{T}$ could be a control point.*

A vector containing the set of control points is notated $\rho = (\rho^{(1)}, \cdots, \rho^{(\tilde{r})})^T$ where $\rho^{(i)} \in \mathcal{T}\ \forall i \in \{1, 2, \cdots, \tilde{r}\}$. The set of control points is combined with a decision rule to define a partition.

Definition 2.3.2. ***Partitioning function:*** *the partitioning function is*

$$\mathcal{P}(\rho, t) = \pi(t)$$

where $\mathcal{P}$ describes a decision rule.

For example, the decision rule could be a Voronoi tessellation [48]. Control points are advantageous because uncertainty in the partition to which an element belongs may be specified in terms of uncertainty in the lower dimensional control points.

2.3.1 Adjusting Mixtures of Random Fields

Modelling with mixtures of random fields provides additional flexibility at the cost of more complex specifications. A field must be specified for every partition. We must also elicit the dependence between the constituent fields of the mixture and how the partitions are defined. Three methods will be discussed in this section: independent partitions, dependent partitions and uncertain partitions.

Independent Partitions

The simplest case is when all partitions are considered mutually independent. Then each partition is modelled as a multivariate random field of the form

$$y_{(k,\omega)}(t) = \mathcal{M}_{(k,\omega)}(t) + \mathcal{W}_{(k,\omega)}(t).$$

Using the approach in Section 2.2.2, the adjusting procedure is applied to each partition separately. There are advantages to using independent segments: inferences are simple, they are computationally more efficient and complications due to mean-non-differentiable fields can be resolved. Independent partitions are also a tractable approach to relaxing the isotropic and stationary residual process assumptions. The residual process for each partition may be given a partition-dependent correlation function with partition-dependent hyper-parameters. However there is a drawback. There is no guarantee that the adjusted expectations are continuous over boundaries; indeed they are unlikely to be.

Dependent Partitions

If an expert believes the mean function is continuous, it is appropriate to specify the partitions as dependent. These beliefs may be incorporated by altering the specification of the component $\mathcal{M}_{(k)}(t)$. The specification is now made using splines (piece-wise polynomials) with knots defined by the control points (see Appendix A.2). Once the design matrix is re-formulated, adjusted beliefs are calculated using the approach presented in Section 2.2.2. The adjusted expectation will be more consistent with prior beliefs and additional uncertainty will be resolved. There are disadvantages: the level of specification is higher and there are fewer computational benefits when compared with independent partitions.

Uncertain Partitions

If a model is specified using partitions, the next logical step is to consider when there is uncertainty pertaining to the partition to which an element belongs. The novel approach presented here elicits reasonable bounds for control points defining a space χ^{ρ}, referred to as the non-implausible space. The approximate adjusted expectation and variance is then calculated using

$$\mathbb{E}_z[y_{(:)}] \approx \frac{1}{m}\sum_{i=1}^{m}\mathbb{E}_z[y_{(:)};\rho^{(:,i)}] \text{ and} \tag{2.3.2}$$

$$\begin{aligned}\mathbb{V}ar_z[y_{(:)}] \approx{}& \frac{1}{m}\sum_{i=1}^{m}\mathbb{V}ar_z[y_{(:)};\rho^{(:,i)}] \\ &+ \frac{1}{m-1}\sum_{i=1}^{m}(\mathbb{E}_z[y_{(:)};\rho^{(:,i)}]-\mathbb{E}_z[y_{(:)}])(\mathbb{E}_z[y_{(:)};\rho^{(:,i)}]-\mathbb{E}_z[y_{(:)}])^T\end{aligned} \tag{2.3.3}$$

where the $\rho^{(:,i)}$s are samples from χ^{ρ}. $\mathbb{E}_z[y_{(:)};\rho^{(:,i)}]$ is the adjusted expectation of y given the specification defined by control point combination $\rho^{(:,i)}$. This approach is effectively making two belief statements: i) all points in the set are *a-priori* equally likely to be the control points and ii) the observation does not change beliefs about the control points.

A refocusing approach is adopted to incorporate that the observations change beliefs about the control points. The technique removes points from χ^{ρ} using a measure, referred to as an implausibility measure, which is designed so that a value over a threshold implies that it is highly unlikely that $\rho^{(:,i)}$ produced the observation. To calculate implausibility, the observation vector z is split into two $\{z(T'), z(T^*)\}$ where $\{T', T^*\} = T$. Beliefs about

$z(T')$ are adjusted using $z(T^*)$ and implausibility calculated using, for example

$$\mathcal{IM}(z,\rho) = \max_{t'\in T', k\in\{1,\cdots,q\}} \left\{ \left| \frac{z_{(k)}(t') - \mathbb{E}_{z(T^*)}[z_{(k)}(t');\rho]}{\sqrt{\mathbb{V}ar[\epsilon^{(2)}_{me(k)}(t')] + \mathbb{V}ar_{z(T^*)}[z_{(k)}(t');\rho]}} \right| \right\}.$$

If $\mathcal{IM}(z,\rho)$ is greater than a threshold, ρ is removed from the set. Adjusted beliefs may be calculated using equations 2.3.2 and 2.3.3 where the $\rho^{(:,i)}$ are sampled from the refocused set.

This approach to prediction is an example of a more general method of Bayesian statistical inversion and prediction which will be a common feature of this thesis. The refocusing constitutes inversion and the averaging over the refocused set prediction. It is efficient, pragmatic and includes the key uncertainties in the problem, while avoiding unreasonable requirements for expert specification and computational effort. A detailed discussion of the method is reserved until Chapter 5 where it plays a more pivotal role.

2.3.2 Toy Example

At this point, a toy example is introduced in order to illustrate adjusting beliefs about multivariate random fields in a low dimensional form. The aim of the section is to exemplify belief adjustment for multivariate random fields. The example was designed to show i) a multivariate random field where the mean-differentiable assumption is not consistent with the field, ii) how the method performs when the specification is inconsistent and iii) how the adjustment performs when the specifications are more consistent. To achieve this purpose, the adjustments will be performed using a number of specifications: i) multivariate random field, ii) a mixture of multivariate random fields with independent partitions, iii) mixture of multivariate random fields with dependent partitions and iv) mixture of multivariate random fields with uncertain partitions. The model takes the form

$$z = M(y) \oplus R^L(y) \oplus \epsilon^G_{me} \oplus \epsilon^L_{me}$$

where the relationship between outputs and the distributions used to sample the data are given by

$$\begin{aligned} y_{(2)}(t)|y_{(1)}(t) &= 0.1 + 1.5y_{(1)}(t) + 0.5(y_{(1)}(t))^2, \\ R^L(y_{(1)}(t)) &\sim \mathcal{N}(0, 0.001), \\ \epsilon_{me}{}^{L}_{(:)} &\sim \mathcal{N}(0, \Sigma_{\epsilon_{me}} \otimes \mathcal{C}), \\ \epsilon_{me}{}^{G}_{(k)}(t) &= (1 + t + t^2)\gamma_{(:,k)} \text{ and} \\ \gamma_{(:)} &\sim \mathcal{N}(0, I_2 \otimes \Sigma_\gamma) \end{aligned} \tag{2.3.4}$$

where

$$\Sigma_{\epsilon_{me}} = \begin{bmatrix} 0.001 & 0 \\ 0 & 0.001 \end{bmatrix}, \Sigma_\gamma = \begin{bmatrix} 1 \times 10^{-5} & 0 & 0 \\ 0 & 1 \times 10^{-7} & 0 \\ 0 & 0 & 1 \times 10^{-10} \end{bmatrix}$$

and $\mathcal{C}$ is the correlation matrix defined by the correlation function given in equation 2.2.8, with p = 1 and correlation length 3. The population mean $M(y)$ for the multivariate random field is shown in Figure 2.2; this is the quantity about which inferences are made. It is mean-non-differentiable and, as such, is challenging to fit with a mean-differentiable field. Figure 2.3 shows a sampled observation from the toy example with a 3 sigma credible interval. This is the vector that will be used to adjust beliefs about $M(y)$.

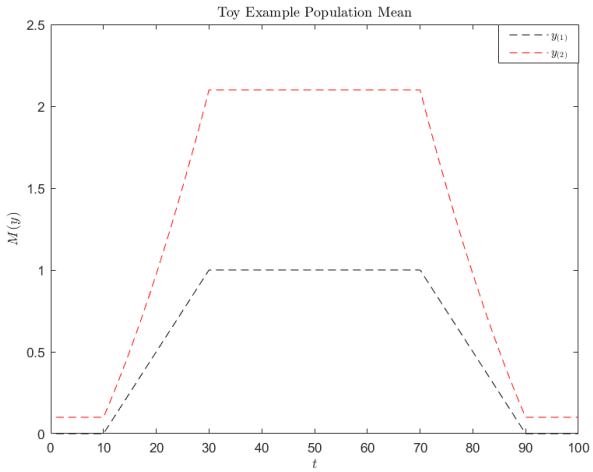

Figure 2.2: The population mean $M(y)$ for both outputs in the toy example. This is the quantity about which inferences will be made.

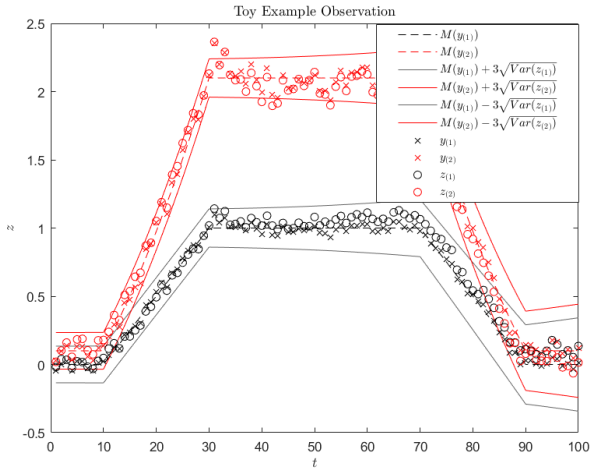

Figure 2.3: Circles: sampled observation z from the toy example. Crosses: sampled y. Dashed lines: population mean $M(y)$. Lines: 3sigma credible intervals.

To illustrate the adjusting procedure for multivariate random fields, beliefs abut the population mean of the toy example will be adjusted with a sampled observation. For all adjustments, the specifications of the components in the set $\{R^L(y), \epsilon^G_{me}, \epsilon^L_{me}\}$ were given by

$$\mathbb{E}[\gamma_{(:)}] = 0,\ \mathbb{E}[\epsilon^L_{me}] = 0 \text{ and}$$

$$\mathbb{V}ar[\gamma_{(:)}] = I_2 \otimes \begin{bmatrix} 1 \times 10^{-5} & 0 & 0 \\ 0 & 1 \times 10^{-7} & 0 \\ 0 & 0 & 1 \times 10^{-10} \end{bmatrix}$$

.

The measurement error covariance structure was specified by

$$\mathbb{C}ov[\epsilon^L_{me(k)}(t), \epsilon^L_{me(k')}(t')] = 0.001\delta_{(k,k')} \exp\left(-0.5\frac{|t-t'|}{3}\right)$$

where $\delta_{(k,k')} = 1$ if $k = k'$ and 0 otherwise. The final specification was $\mathbb{V}ar[R^L(y_{(k)}(t))] = 0.001\ \forall k, t$. This specification was chosen to closely approximate the sampling model used to create the data. This assesses how well belief adjustments for the population mean perform given the other components of the problem are essentially known.

Multivariate Random Field

The first form for the population mean was

$$M(y_{(k)}(t)) = \beta_{(k)} + \mathcal{W}_{(k)}(t)$$

with second order specification

$$\mathbb{E}[\beta_{(k)}] = 1,\ \mathbb{C}ov[\beta_{(k)}, \beta_{(k')}] = 1\delta_{(k,k')}\ \forall k \text{ and}$$
$$\mathbb{C}ov[\mathcal{W}_{(k)}(t), \mathcal{W}_{(k')}(t')] = 0.1c_{(k,k')} \exp\left(-0.5\frac{|t-t'|}{10}\right)$$

where $c_{(k,k')} = 0.8$ if $k = k'$ and 1 otherwise. This is an example of how the method performs when the specifications are inappropriate. Beliefs about the population mean have been specified as a mean-differentiable field when the quantity itself is mean-non-differentiable.

Figure 2.4 shows the adjusted expectations, approximate credible intervals and standardised residuals for the population mean of the toy example. In this case, the adjusted expectation does not produce a mean-differentiable random field. In order to validate, the correlation function had to be non-differentiable. The problem is that the stationary-isotropic assumptions are inappropriate as the mean function is not accounting for all non-stationary behaviour. Figure 2.5 shows results from a leave one out cross validation procedure. The uncertainty is overestimated which is caused by the correlation function being too rough. Figure 2.6 show the cross validation results when the correlation function was squared exponential and correlation length three. There are a number of points which indicate cause for concern as the uncertainty over non-differentiable points is underestimated. Overall the adjusted expectation is a reasonable representation of the expert's uncertainty. However, mixtures of multivariate random fields are a better method to specify beliefs about the population mean field.

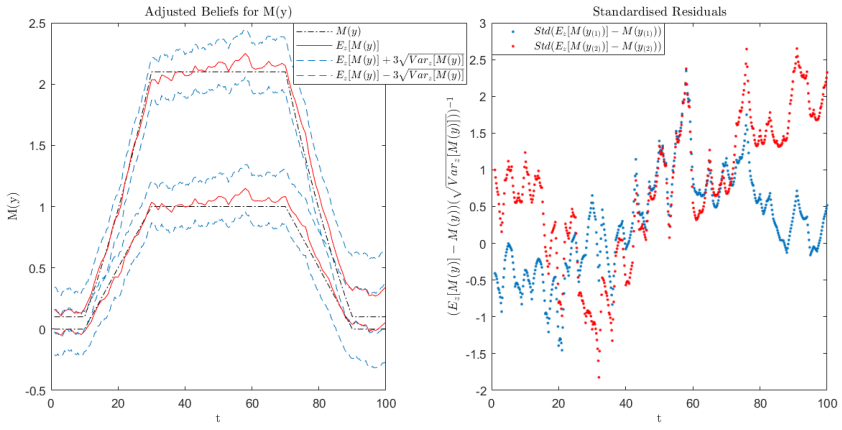

Figure 2.4: Left: adjusted beliefs about the toy example population mean $M(y)$ with belief specification using a mean-differentiable random field. Right: standardised errors for the adjusted version.

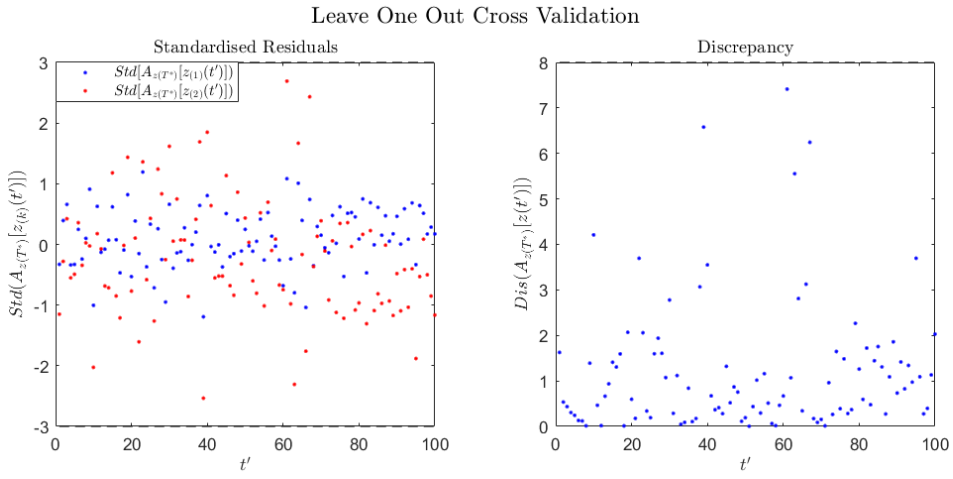

Figure 2.5: Left: standardised residuals of adjusted expectation calculated using leave one out cross validation. Right: discrepancy of adjusted expectation calculated using leave one out cross validation.

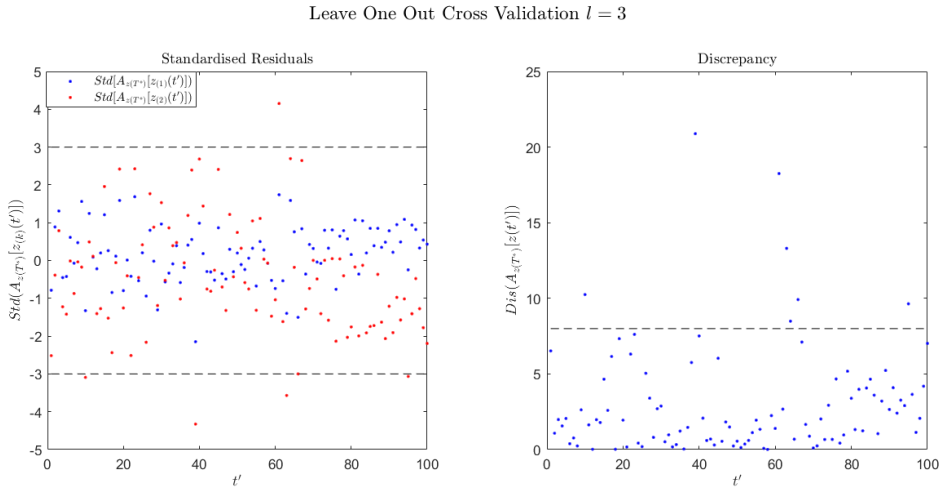

Figure 2.6: Squared exponential correlation function with correlation length three. Left: Standardised residuals of adjusted expectation calculated using leave one out cross validation. Discrepancy of adjusted expectation calculated using leave one out cross validation.

Mixture with known Boundaries and Independent Partitions

The second adjustment is a partition defined by known partitions where the fields belonging to separate partitions are mutually independent. The partition boundaries are defined at the non-differentiable points $(0, 10, 30, 70, 90, 100)$. The specification of the population

mean in partition k is given by

$$M(y_{(k)}(t)) = \sum_{\omega=1}^{5} \pi_{(\omega)}(t) M(y_{(k,\omega)}(t)) \text{ where}$$
$$M(y_{(k,\omega)}(t)) = (1,t)\beta_{(:,k,w)} + \mathcal{W}_{(k,w)}(t).$$

For this example, the second order specification is given by

$$\mathbb{E}[\beta_{(:,k,\omega)}] = (1,0)^T,\ \mathbb{V}ar[\beta_{(1,k,\omega)}] = 1,\ \mathbb{V}ar[\beta_{(2,k,\omega)}] = 2\ \forall k,\omega,$$
$$\mathbb{C}ov[\beta_{(i,k,\omega)}, \beta_{(i',k',\omega')}] = 0\ \forall (i,k,\omega) \neq (i',k',\omega') \text{ and}$$
$$\mathbb{C}ov[\mathcal{W}_{(k,\omega)}(t), \mathcal{W}_{(k',\omega')}(t')] = 0.1 c_{(k,k')} \delta_{(\omega,\omega')} \exp\left(-0.5 \frac{|t-t'|^2}{3^2}\right)$$

where $c_{(k,k')} = 0.8$ if $k = k'$ and 1 otherwise.

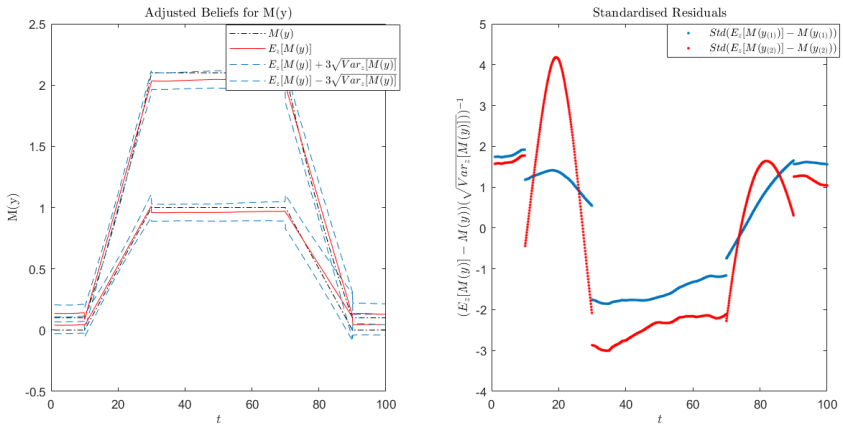

Figure 2.7: Left: adjusted beliefs about the toy example population mean $M(y)$ with belief specification using independent partitions. Right: standardised errors for the adjusted version.

Figure 2.7 shows the toy example where each partition has been adjusted individually. The adjusted expectations are visibly smoother on the plateaus without necessitating significant smoothing over the edges. The uncertainty estimation is reasonable, if slightly underestimated, shown by the standardised errors remaining between $[\pm 3]$. The plot of the standardised residuals contains discontinuities and a large amount of correlation which could be undesirable. This problem can be addressed with dependent partitions or uncertain partitions. Figure 2.8 shows cross validation diagnostics for the independent partitions. Although the adjusted expectations look a better representation of known truth, the diagnostics still have a number of concerning points.

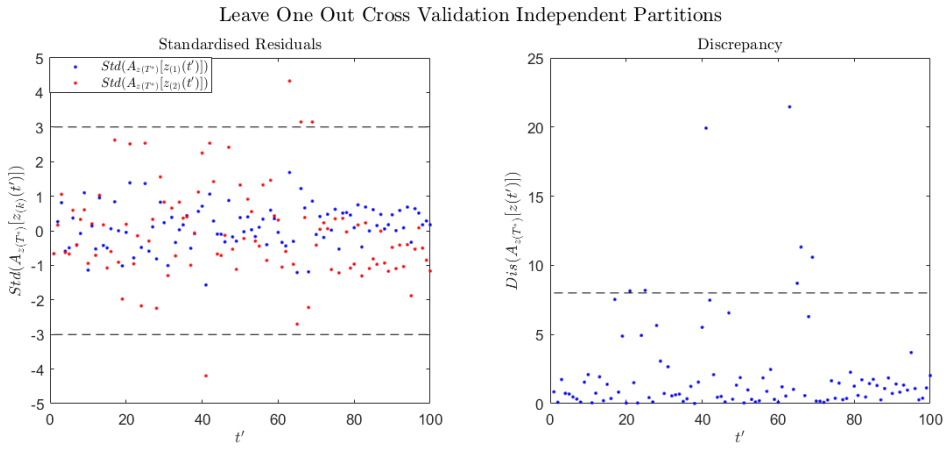

Figure 2.8: Left: standardised residuals of adjusted expectation calculated using leave one out cross validation. Right: discrepancy of adjusted expectation calculated using leave one out cross validation.

Mixture with Known Boundaries and Dependent Partitions

The third specification is where the mean function is specified as continuous but non-differentiable. This is performed by changing the mean function specification to one defined by a first order B-spline (see Appendix A.2). Each segment has a linear model and the set of linear models are tied together at the knots. The second order specifications were given by

$$\mathbb{E}[\beta_{(:,k)}] = (0, 1.5, 1.5, 0)^T, \mathbb{V}ar[\beta_{(:)}] = I_2 \otimes \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ and}$$

$$\mathbb{C}ov[\mathcal{W}_{(k)}(t), \mathcal{W}_{(k')}(t')] = 0.1 c_{(k,k')} \exp\left(-0.5 \frac{(t-t')^2}{3^2}\right).$$

In this case, a single residual process was considered appropriate for the entire field.

Figure 2.9 shows the adjusted beliefs about the population mean where the mean function was specified using a first order B-spline. The right hand plot shows the adjusted regression component is now continuous, the process can be made smooth enough to match the population mean field and additional uncertainty has been resolved. Figure 2.10 shows the cross validation results for dependent partitions. There are a small number of points which would indicate cause for concern but overall it performs well.

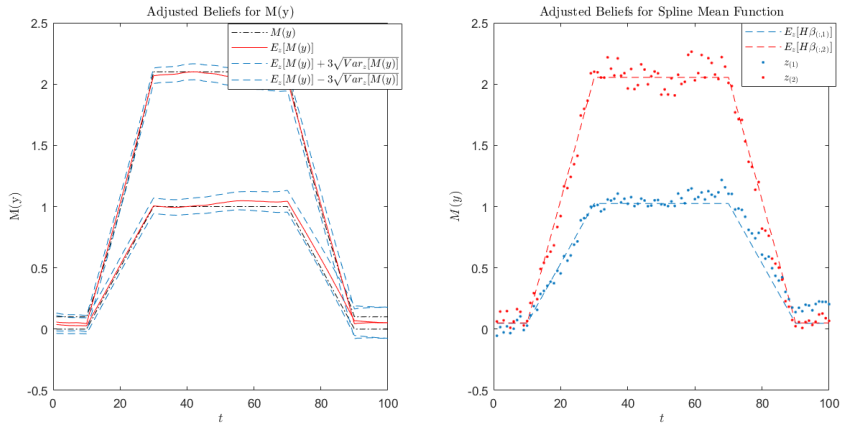

Figure 2.9: Left: adjusted beliefs about the toy example population mean $M(y)$ with belief specification using dependent partitions. Right: adjusted beliefs about the mean function $\mathcal{M}(y)$.

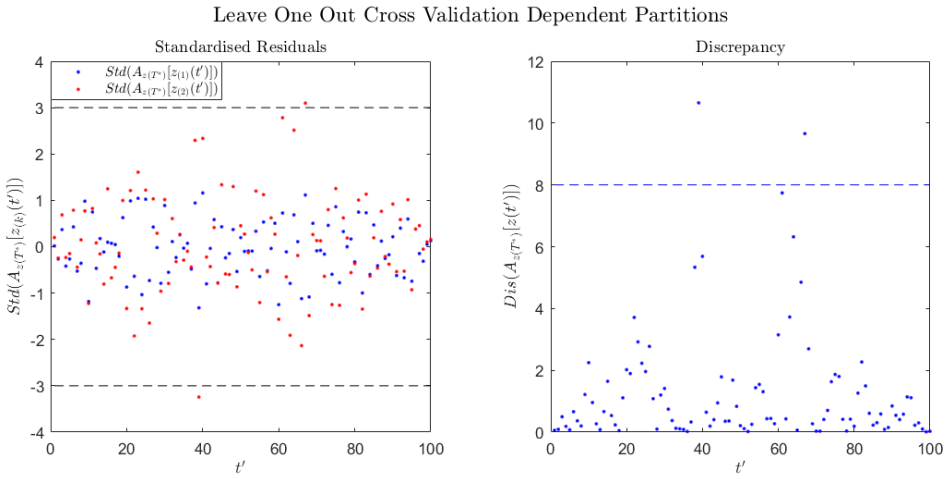

Figure 2.10: Left: standardised residuals of adjusted expectation calculated using leave one out cross validation. Right: discrepancy of adjusted expectation calculated using leave one out cross validation.

Mixture with Uncertain Boundaries and Dependent partitions.

The final adjustment is where the number of partitions is known but there is uncertainty in the control points. To specify the uncertainty in the boundary it was assumed that the boundaries could be within 6 of $(10, 30, 70, 90)$. The set of control points was refocused using half the observation vector, $z(T^*)$, to adjust beliefs about the other half of the observation vector, $z(T')$. Here $T^* = (t^{(1)T}, t^{(3)T}, \cdots t^{(n-1)T})^T$ and $T' = (t^{(2)T}, t^{(4)T}, \cdots t^{(n)T})^T$.

Then if

$$\mathcal{IM}(z,\rho) = \max_{t' \in T', k \in \{1,\cdots,q\}} \left\{ \left| \frac{z_{(k)}(t') - \mathbb{E}_{z(T^*)}[z_{(k)}(t');\rho]}{\sqrt{\mathbb{V}ar[\epsilon^{(2)}_{me_{(k)}}(t')] + \mathbb{V}ar_{z(T^*)}[z_{(k)}(t');\rho]}} \right| \right\} > 2$$

ρ is removed from the set of possible control point combinations. Here Std is defined in equation 2.1.1. Figure 2.11 shows outputs of the refocusing procedure. The black points are the control points which are highly unlikely to have produced the data. The red points are those which there is no evidence to suggest they did not produce the data. The blue histograms are the prior set of control points and the orange histograms are the refocused set of control points.

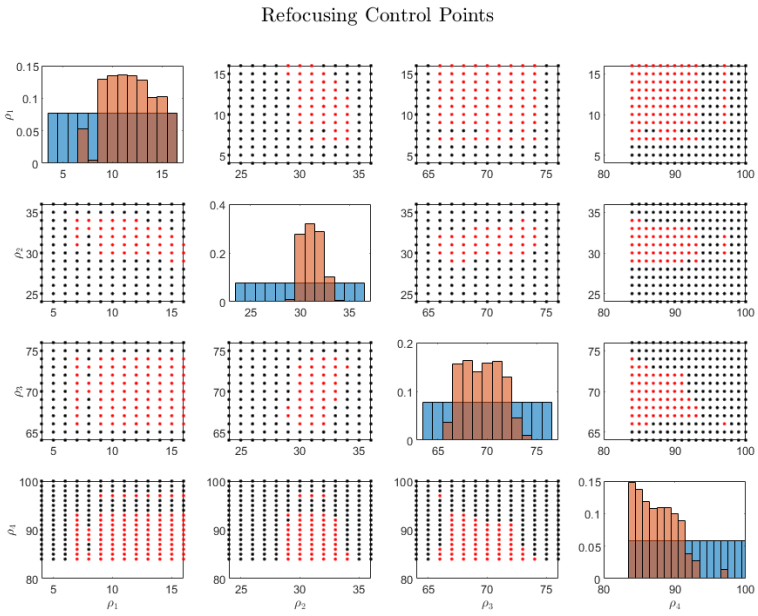

Figure 2.11: Plot illustrating the refocusing procedure. Black points are control point combinations which are ruled out in the refocusing procedure. Red point are those for which are not ruled out. The histograms show the marginal distribution before and after refocusing. Blue is the marginal distribution before refocusing and orange is the marginal distribution after focusing.

A set of 1000 control point combinations were sampled from the refocused set and the adjusted expectation and variance – given the control point combination – computed for all 1000 samples. The second order specification for the field, given a control point combination, is unchanged from the previous specification. The adjusted expectation and variance is then calculated using equation 2.3.2 and 2.3.3. Figure 2.12 shows the adjusted beliefs about the population mean given uncertain boundaries. Making the control points uncertain and averaging over adjusted beliefs, given a control point combination, has a

smoothing effect on the adjusted expectation and increases the adjusted variance. However, if an expert is uncertain about boundaries, this is a better representation of their adjusted beliefs. It also has an advantage when combined with independent partitions in that the averaging process can smooth out the boundaries. This means that the adjusted expectation is mean-differentiable even though the adjusted expectation, given a control point combination, is not.

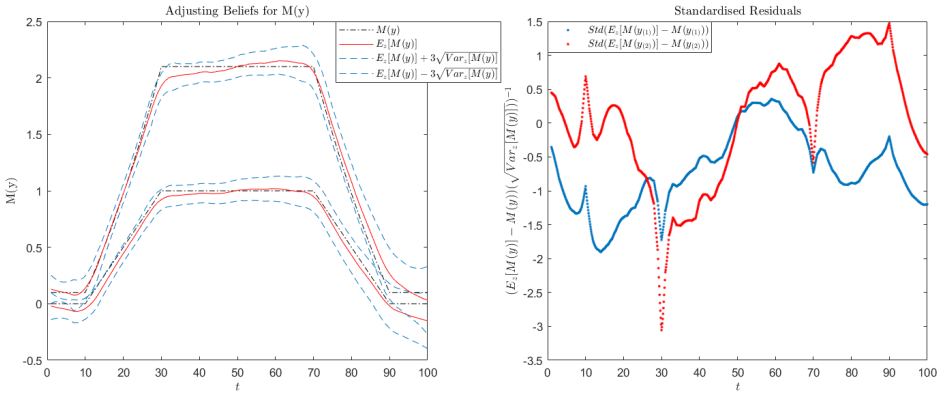

Figure 2.12: Left: adjusted beliefs about the toy example population mean $M(y)$ with uncertain control points. Right: standardised errors for the adjusted version.

The *toolbox* set out in this section provides a flexible methodology for modelling random fields and adjusting beliefs about these fields using vectors of elements on random fields. As more flexibility is obtained, there is a computation and/or specification cost. In our experience, the independent mixtures approach has the best pain-to-gain ratio and it is therefore the recommended starting point. If discontinuities are a problem, it is easier to fix them by adding uncertain control points because averaging over the non-implausible set can produces a continuous adjusted expectation and the other methods scale less well when the dimension of t increases.

2.4 Approximate Kernel Methods

For large observation vectors, the adjusting procedure is still computationally challenging as the size of the matrix to be inverted becomes too large to store in memory and invert. In this section, a novel method to efficiently approximated the analysis is presented. Bayes Linear adjustments may be written as a weighted average of the elements of the observation vector, which implies an equivalent kernel approach may be used. This is only possible if the distances between locations on the y and z fields are the same.

When adjusting the expectation of $y_{(k)}$ by random vector $z_{(k)}$, adjusted expectation may be written as a linear combination of the elements of $z_{(k)}$. For example, if the prior expectation of $z_{(k)}$ is zero, the adjusted expectation of the residual component for a constituent random field may be written as

$$\mathbb{E}_{z_{(k)}}[\mathcal{W}_{(k)}(t')] = \mathbb{C}ov[\mathcal{W}_{(k)}(t'), z_{(k)}](\mathbb{V}ar[z_{(k)}])^{-1} z_{(k)} = K(t') z_{(k)} \tag{2.4.1}$$

where $K(t')$ is a $(1 \times n)$ vector referred to as the adjustment kernel. If $K(t')$ can be assumed to be approximately invariant to t', the adjustment at every location on the field can be performed efficiently with the fast Fourier transform (FFT) and the convolution theorem.

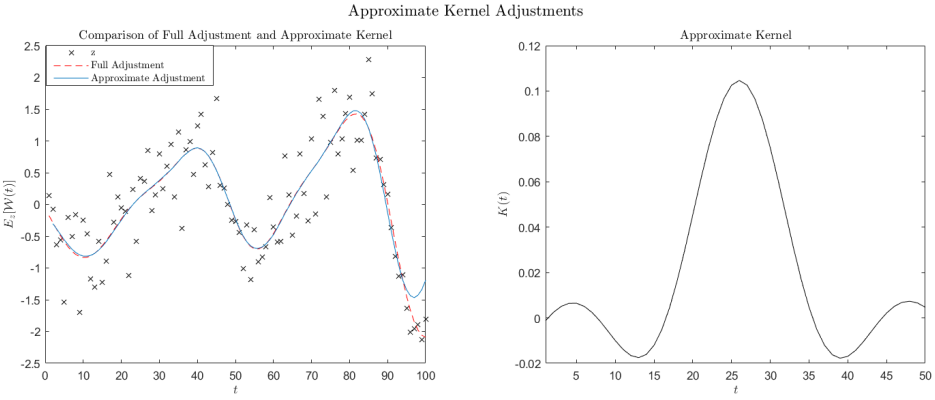

Figure 2.13: Left: adjusted expectations of a stationary random field using the full adjustment and the approximate kernel adjustment. Right: approximate kernel used to calculate the approximate adjusted expectation.

Figure 2.13 shows the full and approximate adjusted beliefs, and the approximate kernel for a stationary process. The difference between adjustments is small except at the edges where the approximation is least valid. The kernel plots are an illustrative representation of how belief adjustment transforms into weights and acts on the data. This is useful when discussing Bayesian statistics with experts from other fields, for example signal processing.

A significant problem is that many fields are not stationary, which means that an approximate kernel cannot be designed to have a computational benefit. The approach taken here i) first adjusts the regression parameters and ii) then adjusts the residual process. Adjusting the regression parameters first can remove a large portion of the non-stationary behaviour. This means that the kernel approach can still approximate the full

adjustment sufficiently well. Assuming $\epsilon_{me}(t)$ is stationary, the new kernel is defined by

$$\begin{aligned}\mathbb{E}_{z_{(k)}}[\mathcal{W}_{(k)}(t')] &= \mathbb{C}ov[\mathcal{W}_{(k)}(t'), z_{(k)}](\mathbb{V}ar[z_{(k)}])^{-1}(z_{(k)} - H\mathbb{E}_{z_{(k)}}[\beta_{(:,k)}]) \\ &\approx \mathbb{C}ov[\mathcal{W}_{(k)}(t'), z^{(k)}](\mathbb{V}ar[\mathcal{W}_{(k)}(T)] + \mathbb{V}ar[\epsilon_{me(k)}(T)])^{-1}(z_{(k)} - H\mathbb{E}_{z_{(k)}}[\beta_{(:,k)}]) \\ &\approx K(t')(z_{(k)} - H\mathbb{E}_{z_{(k)}}[\beta_{(:,k)}]).\end{aligned} \tag{2.4.2}$$

The process will be effective when the majority of the global variation has been resolved so that the remaining variation is approximately stationary. In other words, $H\mathbb{V}ar_{z_{(k)}}[\beta_{(k)}]H^T$ is small compared to $\mathbb{V}ar[\mathcal{W}_{(k)}(T)]$. Adjusting the global variation requires the full data-set. However, the adjustment of the regression parameters of a multivariate random field can often be well approximated using a *small* (under 1000 elements) subset of the data. If this is the case, the adjustment of every location on a random field can be made efficient, which enables real time inference using Bayes linear methods even in high dimensional cases. This process is recommended when efficient inference is essential and the majority of the global variation may be resolved using a *small* subset of the observation vector.

GLS Estimates

If the GLS approximation (key-formula 2.2.1) of adjusted expectations is utilised, the approximate kernel is given by

$$K = \mathrm{Corr}[\mathcal{W}_{(k)}(t'), \mathcal{W}_{(k)}(T)](\mathrm{Corr}[\mathcal{W}_{(k)}(T)] + wI_n))^{-1}. \tag{2.4.3}$$

2.5 Conclusions

In this chapter, a flexible and tractable methodology for specifying and adjusting beliefs about multivariate random fields has been presented. Every subsequent chapter will use the methodology presented in this chapter. The following chapter uses a mixture of multivariate random fields with independent partitions to model X-ray images of chocolate bars. Chapter 4 uses multivariate random fields to encode uncertainty about the output of a complex computer model. Chapter 5 uses multivariate random fields to specify prior beliefs about *infinite* dimensional inverse problems. Chapter 6 uses multivariate random fields to specify beliefs about the difference between a simulator and the system it purports to represent.

In this work, beliefs are adjusted using high resolution X-ray images or simulations of high resolution X-ray images. In both cases, there was a large amount of data available. Therefore, the majority of the heavy lifting is done by the GLS approximation as only strong prior beliefs are likely to have a large impact on the inference. Given that there was a large amount of data available and a key consideration of this work was *real-time* feasibility, separability is often assumed as it makes calculations more efficient and avoids memory bottlenecks.

The next chapter focuses on solving problem statement 1.1.1; finding plastic in chocolate. Multivariate random fields will be used to model X-ray images of chocolate bars, second order exchangeable random fields will be used to make inferences about the populations to which the X-rays images belong, approximate kernel methods enable efficient inference and discrepancy measures will be used to decide if a chocolate bar contains a plastic contaminant.

Chapter 3

Anomaly Detection and Second Order Exchangeability

A chocolate bar manufacturer has a problem with components breaking off the production line and contaminating the product. The contaminants of particular interest are made of polystyrene or polycarbonate with density and chemical composition similar to that of chocolate. When chocolate bars are observed with conventional mono-spectral X-ray examination equipment, the effect of contamination is not sufficiently great to be distinguished from product variation. As a result, these examination techniques and usual anomaly detection methodologies are unable to consistently detect plastic contaminants in chocolate bars. The novel Multi Absorption Plate (MAP [1]) technology [46] provides a multi-spectral X-ray image which enables plastic contamination to be distinguished from product variation.

The requirement of multi-spectral X-ray examination systems to detect contaminants with similar absorption to the product (for example, low density plastics in chocolate) is well documented [11, 12]. The advantage of MAP technology over, for example, dual energy systems is that it is a simple engineering solution which means all the information is available in a single image and multiple acquisitions are not required. As speed is a prominent consideration in quality control systems, this is an advantage. A number of methodologies have been developed for anomaly detection using X-ray imaging (both mono-spectral and multi-spectral). For example, discriminant analysis [15, 16, 18], logistic

[1] This is not to be confused with *maximum-a-posteriori*, which fortunately does not appear in this thesis.

regression [17], Gaussian mixture classification [4], adaptive thresholding [11,14] and neural networks [13, 18, 19]. We attempted to use logistic regression, discriminant analysis and neural networks (using the Matlab [20] Statistics and Machine Learning toolbox) to pixel-wise classify whether a pixel contained plastic contamination. The signal was too weak for consistent pixel-wise identification and the addition post-image-processing and adaptive thresholding (using the Matlab [20] Image Processing toolbox) did not fix the problem. Furthermore, as a training set for plastic was required, the method could also fail to detect contaminants consisting of other materials. We also attempted to use edge detection, filtering and adaptive thresholding (using the Matlab [20] Image Processing toolbox) to identify suspicious regions on which further analysis could be applied. These methods failed to detect regions which contained small plastics, meaning, no matter how accurate subsequent tests were, small plastic contaminants would never be detected.

The method presented in this chapter uses a statistical model of non-anomalous population of chocolate bars, trained using a set of 20 chocolate bars belonging to the non-anomalous population. This extra information (from the training samples and the MAP) combined with Bayes linear methods enables all regions containing plastic to be efficiently identified. By testing all the pixels in regions which are identified as possibly containing contamination together, the problem of a weak pixel-wise signal is overcome. As the method is only trained using non-anomalous chocolate bars, other contaminants (that are more different to chocolate than plastic) can still be accurately detected. Therefore, the method presented overcomes the problems with i) mono-spectral X-ray examination equipment and ii) conventional contaminant detection methodologies with i) the MAP and ii) a spatial model of the non-anomalous population and Bayes linear methods.

Our solution to the plastic in chocolate problem requires a methodology with the following features: a method for efficiently modelling X-ray images, a method for learning about the population to which the X-ray images belong and a method for deciding if a contaminant is present. This chapter will be organised as follows: Section 3.1 will describe the anomaly detection problem, Section 3.2 will introduce second order exchangeability, Section 3.3 will discuss second order exchangeable random fields and Section 3.5 will apply the methodology presented so far to finding plastic contaminants in chocolate bars.

3.1 Anomaly Detection Preliminaries

The use of Bayes linear methods will be exemplified using the general class of problem to which the chocolate bar application belongs. The general class of problem is referred to by many names: anomaly detection, novelty detection, outlier detection or One-Class-Classification [1–3]. Here it will be referred to as anomaly detection. It has been applied to a wide variety of problems: detecting defects in mass produced products [6, 8], distinguishing abnormal growth in tissue [4], identifying suspicious behaviour in shipping lanes [7], recognising malicious software on computers [9] and deducing unusual market behaviour [10].

Anomaly detection is a special case of the more general problem of classification; the process of identifying the class to which an observation belongs. There are three prominent fields which address this problem: traditional statistics, machine learning and Bayesian statistics.

The traditional approach is often based on regressing a class label against an observation. If A is an indicator variable describing whether an observation is in a class or not and z is a vector of observations, inference may be performed using

$$\mathbb{E}[A|z] = \mathcal{L}(h(z)\beta).$$

Here $h(z)$ is a vector of basis functions, β are regression parameters and $\mathcal{L}$ is a function used in some models to ensure $\mathbb{E}[A|z] \in [0, 1]$. For example, Fisher's linear discriminant does not use $\mathcal{L}$ but in logistic regression $\mathcal{L}$ is the logit link function (see [107]). The parameters β are inferred using maximum likelihood and new observations are classified using a threshold. Other approaches include; i) mixtures modelling and ii) decision rules based on Mahalanobis distances [120].

There has also been a large amount of research on classification under the umbrella of machine learning. There are a number of popular approaches: k-means, naive Bayes, decision trees, Gaussian process logistic regression, support vector machines and neural networks [21,30]. Machine learning techniques often have a large number of parameters so perform well when there is a vast amount of data available, and it is useful for solving the problem.[2] They excel at assimilating large data-sets, but are less focused at incorporating

[2] A vast amount of data is not always useful. For example, if the question is whether there is life on any other planet, one instance will do.

relevant prior information [21].

In Bayesian inference, classification is performed by calculating an expert's probability of being in class A given the observation z. This probability is given by Bayes' theorem [98]

$$\mathbb{P}(A=1|z) = \frac{\mathbb{P}(z|A=1)\mathbb{P}(A=1)}{\mathbb{P}(z|A=1)\mathbb{P}(A=1) + \mathbb{P}(z|A=0)\mathbb{P}(A=0)}. \tag{3.1.1}$$

Bayesian inference gives a coherent way of including beliefs about i) relative population sizes and ii) the observation given the class label.

In classification, both $\mathbb{P}(z|A=0)$ and $\mathbb{P}(z|A=1)$ are updated with data. In anomaly detection, only one is. This is usually because the training-set of observations are from the non-anomalous population. This may be because i) there are many possible types of anomalies, ii) anomalies are infrequent or iii) anomalies present themselves in novel forms. A classification approach requires that every class is well represented in the training-set. If this requirement is impracticable, anomaly detection is preferred to classification. However, as classification uses more information, given appropriate data it should always be at least as good as anomaly detection.

The traditional approach to anomaly detection is usually some variant of significance testing. The distribution of the non-anomalous population is inferred from the training-set using, for example, non-parametric density estimation. A test is performed under the null hypothesis that a new observation belongs to the same distribution as the training-set [8]. Machine learning approaches use the error in prediction to make decisions or complement the training-set with synthetic observations in order to apply the usual classification techniques.

An expert may have beliefs about causes of more-likely anomalies and their effect on the observation. For example, their expectation for an observation belonging to the anomalous population will be larger than if it belonged to the non-anomalous population. The Bayesian approach to anomaly detection enables the incorporation of prior beliefs about the i) relative population size, ii) distribution of an observation from the anomalous population and iii) distribution of an observation from the non-anomalous population. It is also coherent with the Bayesian approach to classification; inference is still performed using equation 3.1.1 where the prior specification of $\mathbb{P}(z|A=1)$ or $\mathbb{P}(z|A=0)$ remains unchanged.

The method proposed in this thesis will be set out within the subjective Bayesian framework. The approach follows the same theme as other statistical approaches; beliefs about the population of non-anomalous observations will be adjusted with observations

from that population. Then inference will be performed using the distance from expected, given the observation belongs to the non-anomalous population, by some appropriate metric. The key features of this approach are as follows:

1. Bayes linear methods (discussed in Section 2.1.1) will be used to adjust beliefs,
2. the observations will be modelled as multivariate random fields (discussed in Section 2.2),
3. second order exchangeability (discussed in Section 3.2) will be used to make inferences about the populations to which the observations belong and
4. approximate kernel methods (discussed in Section 2.4) will enable efficient adjustments.

Together these features (Bayes linear methods, multivariate random fields, second order exchangeability and approximate kernels) provide a general methodology to produce an efficient automated anomaly detection algorithm.

3.2 Second Order Exchangeability (SOE)

Given that the methodology is in place for adjusting beliefs about random quantities, the next step is adjusting beliefs about the population to which the random quantities belong. The concept used in this work is exchangeability, which encodes the belief that the entity's label (the superscript i on the $S^{(i)}$) contains no additional information about the entity. Formally, a set of random quantities $\{a^{(1)}, \cdots a^{(N)}\}$) are exchangeable when a belief specification is unaffected by permutation, that is

$$\mathbb{P}(a^{(1)}, \cdots, a^{(N)}) = \mathbb{P}(a^{(\zeta_{(1)})}, \cdots, a^{(\zeta_{(N)})})$$

where the vector ζ is a permutation of $\{1, \cdots, N\}$. Its importance to the subjective Bayesian approach is given by the representation theorem, proposed by de Finetti for Bernoulli random variables in [94]. The theorem implies that, given it is judged that an infinite set of binary random quantities are exchangeable, analysis proceeds as if any subset of this sequence contains independent samples from a Bernoulli distribution with parameter p. This theorem has been extended to general random quantities and states that, given it is asserted that set of random quantities are a subset of an infinite second

order exchangeable sequence, we proceed as if each member of the set are independent samples from a true but unknown distribution. For further details on the fundamental implications of exchangeability on the subjective Bayesian approach, see [51, 99, 112].

The representation theorem is a fundamental theorem which gives a logical reason why many common statistical analyses work. It is, however, an existence theorem and does not state in general the distribution from which the observations are sampled. Therefore, like many full Bayes approaches, it suffers from difficulties in implementation as the level of specification required is too high for an expert to meaningfully make. A simpler formulation is required to be able to use the ideas of exchangeability in large complex problems.

Second Order Exchangeability (SOE)

Second order exchangeability is a simpler formulation which complements the Bayes linear approach.

Definition 3.2.1. ***Second order exchangeable sequence:*** *the set* $\{y_{(:)}^{(1)}, y_{(:)}^{(2)}, \cdots\}$ *is second order exchangeable if the second order belief specification is given by*

$$\mathbb{E}[y_{(:)}^{(i)}] = \mu,\ \mathbb{V}ar[y_{(:)}^{(i)}] = \Sigma\ \forall i \text{ and } \mathbb{C}ov[y_{(:)}^{(i)}, y_{(:)}^{(j)}] = \Gamma\ \forall i \neq j.$$

If a set of random vectors are second order exchangeable, every member of the set may be written in the form

$$y_{(:)}^{(i)} = M(y_{(:)}) \oplus R^{(i)}(y_{(:)}) \tag{3.2.1}$$

with belief specification

$$\mathbb{E}[M(y_{(:)})] = \mu,\ \mathbb{E}[R^{(i)}(y_{(:)})] = 0,\ \mathbb{V}ar[M(y_{(:)})] = \Gamma,$$
$$\mathbb{V}ar[R^{(i)}(y_{(:)})] = \Sigma - \Gamma \text{ and } \mathbb{C}ov[R^{(i)}(y_{(:)}), R^{(j)}(y_{(:)})] = 0\ \forall i \neq j.$$

This is referred to as, "The Representation Theorem for an Infinite Sequence of Second Order Exchangeable Random Vectors" in [28] and was first proposed by Goldstein in [95]. Given that a set of random vectors are considered second order exchangeable, we proceed as if each element of that set is equal to the uncorrelated sum of the population mean and a residual from that population mean. It can then be shown that the sample mean is Bayes linear sufficient for learning about the population mean. This means, given the set

of vectors $D^f = \{y_{(:)}^{(1)}, y_{(:)}^{(2)}, \cdots, y_{(:)}^{(N)}\}$, for any $i > N$

$$\mathbb{E}_{D^f}[y_{(:)}^{(i)}] = \mathbb{E}_{\bar{y}}[M(y_{(:)})] \text{ and}$$
$$\mathbb{V}ar_{D^f}[y_{(:)}^{(i)}] = \mathbb{V}ar_{\bar{y}}[M(y_{(:)})] + \mathbb{V}ar[R^{(i)}(y_{(:)})]$$

where $\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_{(:)}^{(i)}$. The second order specification required for a Bayes linear adjustment is

$$\mathbb{E}[\bar{y}] = \mu, \mathbb{V}ar[\bar{y}] = \Gamma + \frac{1}{N}\mathbb{V}ar[R^{(i)}(y_{(:)})] \text{ and } \mathbb{C}ov[\mathcal{M}(y_{(:)}), \bar{y}] = \Gamma.$$

SOE gives a method for learning about the population mean of a distribution using only second order statements about observables. The specifications are reasonable and the calculations simple.

Variance Learning for Random Vectors

Making inferences about a population's mean is relatively easy – compared to its variance – for a number of reasons: the specifications required are more intuitive, the dimension of the random vectors involved are usually smaller and, due to the uncertainty in the mean, the quantity required for a second order exchangeable adjustment of the population variance is not observable. To make inferences about the population variance, the quantity of interest is

$$V^{(i)} = R^{(i)}(y_{(:)})R^{(i)}(y_{(:)})^T.$$

Assuming the sequence $\{V^{(1)}, V^{(2)}, \cdots\}$ is second order exchangeable implies that

$$V^{(i)} = M(V) \oplus R^{(i)}(V)$$

where $M(V)$ is referred to as the population variance. In this way, a SOE representation has been constructed for learning about the population variance. The associated second order specification is given by

$$\mathbb{E}[M(V)] = \Phi = \Sigma - \Gamma,\ \mathbb{E}[R^{(i)}(V)] = 0,\ \mathbb{V}ar[M(V_{(:)})] = V_{M(V)},$$
$$\mathbb{V}ar[R^{(i)}(V_{(:)})] = V_{R(V)} \text{ and } \mathbb{C}ov[R^{(i)}(V_{(:)}), R^{(j)}(V)_{(:)}] = 0\ \forall i \neq j.$$

The problem is that $M(y)$ is unknown in many applications. Therefore, the quantity required is not directly observable and an informative statistic must be used instead. The usual choice is

$$\hat{\Sigma} = \frac{1}{N-1}\sum_{i=1}^{N}(y_{(:)}^{(i)} - \bar{y})(y_{(:)}^{(i)} - \bar{y})^T = M(V) \oplus T.$$

Given fourth order uncorrelated properties

$$\mathbb{C}ov[M(V), R^{(k)}(y_{(:)})R^{(j)}(y_{(:)})^T] = \mathbb{C}ov[R^{(k)}(y_{(:)})R^{(j)}(y_{(:)})^T, R^{(i)}(V_{(:)})] = 0$$

$\forall k \neq j.\ j \neq i$ and

$$\mathbb{C}ov[R^{(k)}(y_{(:)})R^{(j)}(y_{(:)})^T, R^{(w)}(y_{(:)})R^{(u)}(y_{(:)})^T] = 0$$

$\forall\ k > j,\ w > u$ and $k \neq w,\ j \neq u$

have been asserted, the second order specification for $\hat{\Sigma}$ is given by

$$\mathbb{E}[\hat{\Sigma}] = \Phi,\ \mathbb{V}ar[\hat{\Sigma}_{(:)}] = V_{M(V)} + V_T \text{ and } \mathbb{C}ov[M(V_{(:)}), \hat{\Sigma}_{(:)}] = V_{M(V)}.$$

Fourth order uncorrelated properties are implied when exchangeability is assumed. However, as only second order exchangeability has been asserted, this additional assumption is required. The calculations are relatively simple but the specification is difficult. The quantity $V_{M(V)}$ is the uncertainty in the population variance. V_T is the uncertainty in $\hat{\Sigma}$, analogous to the squared standard error of the usual unbiased population variance estimator. It depends on the shape of the distribution of y, most pertinently its kurtosis. In the uni-variate case, the specifications may be made using [28] Chapter 8. In the multivariate case, these specifications are challenging to elicit in a meaningful manner. For example, the $\mathbb{C}ov[M(V_{(i,i)}), M(V_{(i,j)})]$ requires elicitation of the linear relationship between population variance and population covariance. For these reasons, an automated approach requiring more reasonable second order specifications is of interest, see Appendix A.3 for details.

3.3 Second Order Exchangeable Random Fields (SOERFs)

The next problem that will be addressed is extending the second order exchangeability representation to make inferences about the populations to which multivariate random fields belong. This approach was originally developed for second order exchangeable functions, see [42]. If $y^{(i)}$ is a multivariate random field, $\{y^{(1)}, \cdots, y^{(N)}\}$ is a sequence of SOERFs if

1. the sequence $\{y^{(1)}(t), \cdots, y^{(N)}(t)\}$ is second order exchangeable and
2. the covariance between any two locations on two observations is specified by

$$\begin{aligned}\mathbb{C}ov[y^{(i)}(t), y^{(i')}(t')] &= \Sigma(t,t') \text{ for } i = i' \\ &= \Gamma(t,t') \text{ for } i \neq i'.\end{aligned}$$

Given it is assumed that the set is a subset of an infinite sequence, a multivariate random field may be represented as the independent sum of a population mean and residual from the population mean

$$y^{(i)}(t) = M(y(t)) \oplus R^{(i)}(y(t)).$$

The second order specification of the population mean component is given by

$$\mathbb{E}[M(y(t))] = \mu(t),\ \mathbb{C}ov[M(y(t)), M(y(t'))] = \Gamma(t,t')$$

and the residual field has second order specification

$$\begin{aligned}\mathbb{E}[R^{(i)}(y(t))] &= 0,\\ \mathbb{C}ov[R^{(i)}(y(t)), R^{(i)}(y(t'))] &= \Phi(t,t') = \Sigma(t,t') - \Gamma(t,t') \text{ and}\\ \mathbb{C}ov[M(y(t)), R^{(i)}(y(t'))] &= \mathbb{C}ov[R^{(i)}(y(t')), R^{(i')}(y(t))] = 0.\end{aligned}$$

This is the extension to the representation theorem for SOERFs. Given every member of an infinite set $\{y^{(1)}, \cdots\}$ of multivariate random fields is specified in the form given in equation 2.2.6, the regression parameters are second order exchangeable and the residual components are SOERFs. This implies that the components may be written in the form

$$\begin{aligned}\beta^{(i)}_{(:,k)} &= M(\beta_{(:,k)}) \oplus R^{(i)}(\beta_{(:,k)}) \text{ and}\\ \mathcal{W}^{(i)}_{(k)}(t) &= M(\mathcal{W}_{(k)}(t)) \oplus R^{(i)}(\mathcal{W}_{(k)}(t)).\end{aligned}$$

The population mean of the entire field may then be written in terms of the population mean of the global and local components. The residual from the population mean may also be written in terms of local and global components. This is given by two equations

$$\begin{aligned}M(y_{(k)}(t)) &= h(t)M(\beta_{(:,k)}) \oplus M(\mathcal{W}_{(k)}(t)) \text{ and}\\ R^{(i)}(y_{(k)}(t)) &= h(t)R^{(i)}(\beta_{(:,k)}) \oplus R^{(i)}(\mathcal{W}_{(k)}(t)).\end{aligned}$$

Mean Learning for SOERFs

In this section, a novel closed form method for adjusting beliefs about the population mean field is presented. Adjusting beliefs about the population mean of the regression parameters follows from second order exchangeable adjustments of random vectors. It is assumed all fields are adjusted by a vector with nq elements. The sample mean of adjusted expectations

$$\bar{\beta}_{(:,k)} = \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}_{z^{(k)}}[\beta^{(i)}_{(:,k)}] = \frac{1}{N}\sum_{i=1}^{N} M(\beta_{(:,k)}) + R^{(i)}(\beta_{(:,k)}) - \mathbb{A}_{z^{(i)}}[\beta^{(i)}_{(:,k)}]$$

is Bayes linear sufficient for learning about the population mean of the regression parameters. To calculate the second order specification, note that $\{\mathbb{E}_{z^{(1)}}[\beta^{(1)}_{(:,k)}], \cdots, \mathbb{E}_{z^{(N)}}[\beta^{(N)}_{(:,k)}]\}$ is second order exchangeable as

$$\begin{aligned}
\mathbb{E}[\mathbb{E}_{z^{(i)}}[\beta^{(i)}_{(:,k)}]] &= \mathbb{E}[M(\beta_{(:,k)})], \\
\mathbb{V}ar[\mathbb{E}_{z^{(i)}}[\beta^{(i)}_{(:,k)}]] &= \mathbb{R}\mathbb{V}ar_{z^{(i)}}[\beta^{(i)}_{(:,k)}] \ \forall i \text{ and} \\
\mathbb{C}ov[\mathbb{E}_{z^{(i)}}[\beta^{(i)}_{(:,k)}], \mathbb{E}_{z^{(j)}}[\beta^{(j)}_{(:,k)}]] &= K_k \mathbb{V}ar[M(y_{(k)})] K_k^T \ \forall i \neq j \text{ where} \\
K_k &= (\mathbb{V}ar[\beta^{(i)}_{(:,k)}]H^T)(\mathbb{V}ar[z^{(i)}_{(:,k)}])^{-1}.
\end{aligned} \tag{3.3.1}$$

Then as

$$\begin{aligned}
\bar{\beta}_{(:,k)} &= M(\beta_{(:,k)}) + \frac{1}{N}\sum_{i=1}^{N} R^{(i)}(\beta_{(:,k)}) - \mathbb{A}_{z^{(i)}}[\beta^{(i)}_{(:,k)}] \\
&= M(\beta_{(:,k)}) + \bar{R}(\beta_{(:,k)}) - \bar{\mathbb{A}}_z[\beta_{(:,k)}]
\end{aligned}$$

where $\bar{R}(\beta_{(:,k)}) = \frac{1}{N}\sum_i R^{(i)}(\beta_{(:,k)})$ and $\bar{\mathbb{A}}_z[\beta_{(:,k)}] = \frac{1}{N}\sum_i \mathbb{A}_{z^{(i)}}[\beta^{(i)}_{(:,k)}]$. The covariance between $M(\beta_{(:,k)})$ and $\bar{\beta}_{(:,k)}$ is given by

$$\begin{aligned}
\mathbb{C}ov[M(\beta_{(:,k)}), \bar{\beta}_{(:,k)}] &= \mathbb{V}ar[M(\beta_{(:,k)})] - \mathbb{C}ov[M(\beta_{(:,k)}), \mathbb{A}_{z^{(i)}}[\beta^{(i)}_{(:,k)}]] \\
&= \mathbb{V}ar[M(\beta_{(:,k)})] - \mathbb{V}ar[M(\beta_{(:,k)})] + (\mathbb{V}ar[M(\beta_{(:,k)})]H^T)K_k^T.
\end{aligned}$$

Therefore, the second order specification required for the adjustment is given by

$$\begin{aligned}
\mathbb{E}[\bar{\beta}_{(:,k)}] &= \mathbb{E}[\beta^{(i)}_{(:,k)}], \\
\mathbb{V}ar[\bar{\beta}_{(:,k)}] &= K_k \mathbb{V}ar[M(y_{(k)})] K_k^T \\
&\quad + \frac{1}{N}(\mathbb{R}\mathbb{V}ar_{z^{(i)}}[\beta^{(i)}_{(:,k)}] - K_k \mathbb{V}ar[M(y)] K_k^T) \text{ and} \\
\mathbb{C}ov[M(\beta_{(:,k)}), \bar{\beta}_{(:,k)}] &= (\mathbb{V}ar[M(\beta_{(:,k)})]H^T)K_k^T.
\end{aligned}$$

The sample mean of adjusted expectations of the residual processes

$$\bar{\mathcal{W}}_{(k)}(t) = \frac{1}{N}\sum_{k=1}^{N} \mathbb{E}_{z^{(i)}}[\mathcal{W}^{(i)}_{(k)}(t)] = \frac{1}{N}\sum_{i=1}^{N} M(\mathcal{W}_{(k)}(t)) + R^{(i)}(\mathcal{W}_{(k)}(t)) - \mathbb{A}_{z^{(i)}}[\mathcal{W}^{(i)}_{(k)}(t)]$$

is used to adjust beliefs about the residual process. Using similar calculations to the regression parameters, the second order specification is given by

$$\begin{aligned}
\mathbb{E}[\bar{\mathcal{W}}_{(k)}(T)] &= \mathbb{E}[\mathcal{W}^{(i)}_{(k)}(T)] = \mathbb{E}[M(\mathcal{W}_{(k)}(T))], \\
\mathbb{V}ar[\bar{\mathcal{W}}_{(k)}(T)] &= K'_k(\mathbb{V}ar[M(y_{(k)})])K_k'^T \\
&\quad + \frac{1}{N}(\mathbb{R}\mathbb{V}ar_{z^{(k)}}[\mathcal{W}^{(i)}_{(k)}(T)]] - K'_k \mathbb{V}ar[M(y_{(k)})] K_k'^T), \\
\mathbb{C}ov[M(\mathcal{W}_{(k)}(T)), \bar{\mathcal{W}}_{(k)}(T)] &= \mathbb{V}ar[M(\mathcal{W}_{(k)}(T))]K_k'^T \\
&\text{where } K'_k = (\mathbb{V}ar[\mathcal{W}^{(i)}_{(k)}(T)])(\mathbb{V}ar[z^{(i)}_{(k)}])^{-1}.
\end{aligned}$$

Learning about the population mean field is analogous to learning about a random field with a vector, the elements of which are locations on a random field.

GLS Approximation

If GLS estimates are used instead of a full Bayes linear adjustment, the specifications are simpler in that the assumption

$$\mathbb{C}ov[\beta^{(i)}_{(:,k)}, \mathbb{A}[\beta^{*(i)}_{(:,k)}]] = 0 \text{ and } \mathbb{C}ov[\mathcal{W}^{(i)}_{(k)}(t), \mathbb{A}[\mathcal{W}^{*(i)}_{(k)}(T)]] \approx 0$$

where $\mathbb{A}[\mathcal{W}^{*(i)}_{(k)}(T)] = \mathcal{W}^{(i)}_{(k)}(T) - \mathcal{W}^{*(i)}_{(k)}(T)$ and $\mathbb{A}[\beta^{*(i)}_{(:,k)}] = \beta^{(i)}_{(:,k)} - \beta^{*(i)}_{(:,k)}$, are reasonable approximations. When $\mathbb{V}ar[\epsilon_{me(:)}] = 0$, the equality holds. Here $\mathcal{W}^{*(i)}_{(k)}(t)$ and $\beta^{*(i)}_{(:,k)}$ are approximated adjusted beliefs with GLS point estimates, see key formula 2.2.1. The sets $\{\beta^{*(i)}_{(:,k)}, \cdots\}$ and $\{\mathcal{W}^{*(i)}_{(k)}(t), \cdots\}$ are approximately second order exchangeable and second order exchangeable random fields respectively. The second order specification is then approximated by

$$\begin{aligned}
\mathbb{E}[\bar{\beta}_{(:,k)}] &\approx \mathbb{E}[M(\beta_{(:,k)})] \\
\mathbb{V}ar[\bar{\beta}_{(:,k)}] &\approx \mathbb{V}ar[M(\beta_{(:,k)})] + \frac{1}{N}\left(\mathbb{V}ar[\mathbb{A}[\beta^{*(i)}_{(:,k)}]] + \mathbb{V}ar[R^{(i)}(\beta_{(:,k)})]\right) \\
\mathbb{C}ov[M(\beta_{(:,k)}), \bar{\beta}_{(:,k)}] &\approx \mathbb{V}ar[M(\beta_{(:,k)})] \\
\mathbb{E}[\bar{\mathcal{W}}_{(k)}(T)] &\approx \mathbb{E}[M(\mathcal{W}_{(k)}(T)] \\
\mathbb{V}ar[\bar{\mathcal{W}}_{(k)}(T)] &\approx \mathbb{V}ar[M(\mathcal{W}_{(k)}(T)] + \frac{1}{N}\left(\mathbb{V}ar[\mathbb{A}[\mathcal{W}^{*(i)}_{(k)}(T)]] + \mathbb{V}ar[R^{(i)}(\mathcal{W}_{(k)}(T))]\right) \\
\mathbb{C}ov[M(\mathcal{W}_{(k)}(T), \bar{\mathcal{W}}_{(k)}(T))] &\approx \mathbb{V}ar[M(\mathcal{W}_{(k)}(T)].
\end{aligned} \tag{3.3.2}$$

Given that GLS estimates are used to approximate adjusted beliefs in the chocolate bar example, this approximation is employed for population learning for that example as well. Figure 3.1 illustrates adjusting beliefs about the population mean field for the toy example. 20 samples were taken from the model and used to adjust beliefs about the population mean of the global and local component. The second order specification of beliefs about each sample was given by the specification for dependent partitions and known boundaries in Section 2.3.2.

Variance Learning For SOERFs

Variance learning for SOERFs is difficult as the size of the vectors involved are very large and the level of specification challenging. Here the focus will be on learning about the population covariance matrix at every location. The population covariance between two

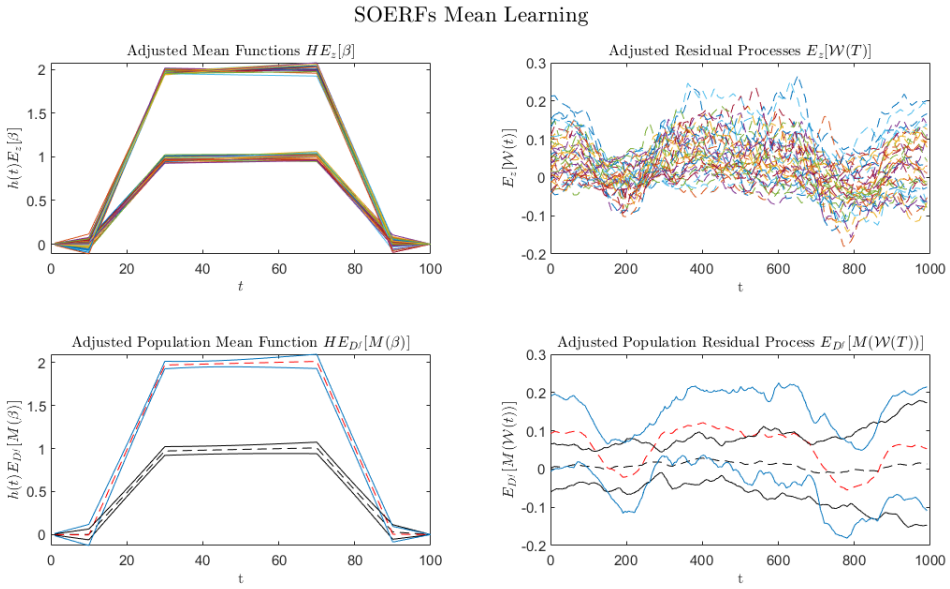

Figure 3.1: Top left: 20 adjusted expectations of the mean function for 20 samples from the toy example. Top right: 20 adjusted expectations of the residual process for 20 samples from the toy example. Bottom left: adjusted beliefs about the population mean of the mean function (red is $H\mathbb{E}_{D^f}[M(\beta_{(:,1)})]$ and black is $H\mathbb{E}_{D^f}[M(\beta_{(:,2)})]$). Bottom right: adjusted beliefs about the population variance of the mean function (red is $\mathbb{E}_{D^f}[M(M(\mathcal{W}_{(1)}(T)))]$ and black is $\mathbb{E}_{D^f}[M(M(\mathcal{W}_{(2)}(T)))]$). The error bars are a 2sigma credible interval.

locations will not be considered. In this approach, the population variance and correlation fields are adjusted separately. The population variance field is notated $M(V_{\mathcal{W}(k)}(t))$ and correlation field $M(C_{(k,k')}(t))$. Then to get the covariance matrix of the residual process at a location, the adjusted expectations for the variances and correlations at that location are combined in the usual way. The matrices

$$V^{(i)}_{\beta(:,:,k)} = R^{(i)}(\beta_{(:,k)})R^{(i)}(\beta_{(:,k)})^T \text{ and } V^{(i)}_{\mathcal{W}(k)}(t) = R^{(i)}(\mathcal{W}_{(k)}(t))R^{(i)}(\mathcal{W}_{(k)}(t))^T$$

are defined, where $V^{(i)}_{\beta(:,:,k)}$ is a $(r \times r)$ and $V^{(i)}_{\beta}$ is a $(r \times r \times q)$ matrix. Then if the sets $\{V^{(1)}_{\beta(:,k)}, \cdots, V^{(N)}_{\beta(:,k)}\}$ (see Appendix A.1 equation 1.1.5 for vectorising three dimensional matrices) and $\{V^{(1)}_{\mathcal{W}(k)}(t), \cdots, V^{(N)}_{\mathcal{W}(k)}(t)\}$ are assumed to be second order exchangeable and second order exchangeable random fields respectively, we may write

$$V^{(i)}_{\beta(:,:,k)} = M(V_{\beta(:,:,k)}) \oplus R^{(i)}(V_{\beta(:,:,k)}) \text{ and } V^{(i)}_{\mathcal{W}(k)}(t) = M(V_{\mathcal{W}(k)}(t)) \oplus R^{(i)}(V_{\mathcal{W}(k)}(t)).$$

Therefore, the quantity of interest is given by

$$h(t)M(V_{\beta(:,:,k)})h(t)^T \oplus M(V_{\mathcal{W}(k)}(t)).$$

This splits the problem up into learning about the population variance of the local and global components. This is similar to variance learning for vectors except both components, for example $\beta^{(i)}_{(:,k)}$ and $M(\beta_{(:,k)})$, are now uncertain. The informative quantities

$$\hat{\Sigma}_{\beta(:,:,k)} = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbb{E}_{z^{(i)}}[\beta^{(i)}_{(:,k)}] - \bar{\beta}_{(:,k)})(\mathbb{E}_{z^{(i)}}[\beta^{(i)}_{(:,k)}] - \bar{\beta}_{(:,k)})^T$$

and

$$\hat{\Sigma}_{\mathcal{W}(k)}(t) = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbb{E}_{z^{(i)}}[\mathcal{W}^{(i)}_{(k)}(t)] - \bar{\mathcal{W}}_{(k)}(t))(\mathbb{E}_{z^{(i)}}[\mathcal{W}^{(i)}_{(k)}(t)] - \bar{\mathcal{W}}_{(k)}(t))^T$$

may be used to adjust beliefs. These statistics may be rearranged to give

$$\hat{\Sigma}_{\beta_{(:,:,k)}} = M(V_{\beta(:,:,k)}) + T_{\beta(:,:,k)} \text{ and } \hat{\Sigma}_{\mathcal{W}(k)}(t) = M(V_{\mathcal{W}(k)}(t)) + T_{\hat{\Sigma}_{\mathcal{W}}(k)}(t). \tag{3.3.3}$$

This is analogous to variance learning for random vectors. The quantity required for a direct second order exchangeable update is not available so an informative statistic is required. This statistic may be decomposed into the population variance plus a residual from the population variance. The structure is significantly more complex as the estimators are biased and the covariances non-zero; for example, $\mathbb{C}ov[M(V_{\beta(:,k)}), T_{\beta(:,k)}] \neq 0$. The other difference is that $M(V_{\mathcal{W}})$ is a multivariate random field, beliefs for which can be specified using covariance functions. For this reason, learning about the population variance field is analogous to adjusting beliefs about a multivariate random field with a vector the elements of which are locations on a multivariate random field. The expectation of the estimators can be calculated from specifications which have already been made. For example, $T_{\beta(:,k)}$ may be decomposed as

$$\begin{aligned}
T_{\beta_{(:,k)}} &= T + T' + T'' \text{ where} \\
T &= \frac{1}{N}\sum_{i=1}^{N} R^{(i)}(M(V_{\beta(:,k)})) - \frac{2}{N(N-1)}\sum_{i<j} R^{(i)}(\beta_{(:,k)})(R^{(j)}(\beta_{(:,k)}))^T \\
T' &= \frac{1}{N}\sum_{i=1}^{N}(\mathbb{A}_{z^{(i)}}[\beta^{(i)}_{(:,k)}])(\mathbb{A}_{z^{(i)}}[\beta^{(i)}_{(:,k)}])^T - \frac{2}{N(N-1)}\sum_{i<j}\mathbb{A}_{z^{(i)}}[\beta^{(i)}_{(:,k)}](\mathbb{A}_{z^{(j)}}[\beta^{(j)}_{(:,k)}])^T \text{ and} \\
T'' &= +\frac{1}{N-1}\sum_{i=1}^{N}(R^{(i)}(\beta_{(:,k)}) - \bar{R}(\beta_{(:,k)}))(\bar{\mathbb{A}}_z[\beta_{(:,k)}] - \mathbb{A}_{z^{(i)}}[\beta^{(i)}_{(:,k)}])^T \\
&\quad + \frac{1}{N-1}\sum_{i=1}^{N}(\bar{\mathbb{A}}_z[\beta_{(:,k)}] - \mathbb{A}_{z^{(i)}}[\beta^{(i)}_{(:,k)}])(R^{(i)}(\beta_{(:,k)}) - \bar{R}(\beta_{(:,k)}))^T
\end{aligned}$$

where $\bar{R}(\beta_{(:,k)}) = \frac{1}{N}\sum_i R^{(i)}(\beta_{(:,k)})$ and $\bar{\mathbb{A}}_z[\beta_{(:,k)}] = \frac{1}{N}\sum_i \mathbb{A}_{z^{(i)}}[\beta^{(i)}_{(:,k)}]$. Given

$$\mathbb{C}ov[R^{(i)}(\beta_{(:,k)}), \mathbb{A}_{z^{(j)}}[\beta^{(j)}_{(:,k)}]] = \mathbb{C}ov[R^{(i)}(\beta_{(:,k)}), R^{(j)}(\beta_{(:,k)})] = 0,$$

the prior expectation is given by

$$\begin{aligned}\mathbb{E}[T] &= 0\\ \mathbb{E}[T'] &= \mathbb{V}ar_{z^{(i)}}[\beta_{(:,k)}] - \mathbb{V}ar[M(\beta_{(:,k)})] + K_k\mathbb{V}ar[M(y_{(k)})]K_k^T\\ \mathbb{E}[T''] &= -2(\mathbb{E}[M(V_{\beta(:,:,k)})] - (H\mathbb{E}[M(V_{\beta(:,:,k)}])^T K_k^T))\end{aligned}$$

where K_k is defined in equation 3.3.1. The variance and covariance specifications for $\{M(V_{\beta_{(:,k)}}), T_{(:)}, T'_{(:)}, T''_{(:)}\}$ are challenging even by the high standard of learning about population variances. There are however a number options if they cannot be made directly: i) use GLS estimates, ii) use a sampling algorithm, iii) use a direct update approach and iv) adjust the population mean and variance of the adjusted expectations. As the GLS approximation method is used in the chocolate bar application, it will be presented below. For the other approaches tested, see Appendix A.4.

GLS Estimates

If the effect prior beliefs on adjusted beliefs is *not large*, GLS estimates may be used instead of a full adjustment. This makes the specifications easier as the approximation

$$\begin{aligned}\mathbb{C}ov[M(V_{\beta(:)}), T_{\beta(:)}] = 0,\ \mathbb{C}ov[M(V_{\mathcal{W}(k)}(t)), T_{\hat{\Sigma}_{\mathcal{W}(k)}}(t)] = 0,\\ \mathbb{E}[T_{\hat{\Sigma}_{\mathcal{W}(k)}}(t)] = \mathbb{V}ar[\mathbb{A}[\mathcal{W}^*_{(k)}(t)]] \text{ and } \mathbb{E}[T_{\hat{\Sigma}_\beta}] = \mathbb{V}ar[\mathbb{A}[\beta^*_{(:)}]]\end{aligned}$$

is more reasonable. Then, assuming $\mathbb{C}ov[T_{\hat{\Sigma}_{\mathcal{W}(k)}}(t), T_{\hat{\Sigma}_{\mathcal{W}(k)}}(t')] = 0$, the adjustments follow using the specifications in [28] chapter 8 and the adjustments of multivariate random fields in Section 2.2.2. This approach is simple and effective when the number of elements in the observation vector is large. The approximated adjusted beliefs are given by

$$\begin{aligned}\mathbb{E}[\hat{\Sigma}_{\beta_{(:,:,k)}}] &\approx \mathbb{E}[M(V_{\beta(:,:,k)})] + \mathbb{V}ar[\mathbb{A}(\beta^*_{(:,k)})],\\ \mathbb{E}[\hat{\Sigma}_{\mathcal{W}(k)}(t)] &\approx \mathbb{E}[M(V_{\mathcal{W}(k)}(t))] + \mathbb{V}ar[\mathbb{A}[\mathcal{W}^*_{(k)}(t)]],\\ \mathbb{V}ar[\hat{\Sigma}_{\beta_{(:,k)}}] &\approx \mathbb{V}ar[M(V_{\beta(:,k)})] + \mathbb{V}ar[T_{\beta(:,k)}],\\ \mathbb{V}ar[\hat{\Sigma}_{\mathcal{W}(k)}(t)] &\approx \mathbb{V}ar[M(V_{\mathcal{W}(k)}(t))] + \mathbb{V}ar[T_{\hat{\Sigma}_{\mathcal{W}(k)}}(t)],\\ \mathbb{C}ov[\hat{\Sigma}_{\beta_{(:,k)}}, M(V_{\beta(:,k)})] &\approx \mathbb{V}ar[M(V_{\beta(:,k)})],\\ \mathbb{C}ov[\hat{\Sigma}_{\mathcal{W}(k)}(t), M(V_{\mathcal{W}(k)}(t))] &\approx \mathbb{V}ar[M(V_{\mathcal{W}(k)}(t))].\end{aligned} \tag{3.3.4}$$

In summary, variance learning is hard and for multivariate random fields harder. The direct specifications are sufficiently challenging that it is unreasonable to expect an expert to be able to make them directly. In the approach used in this thesis, GLS approximations

are used and, given the size of the observation vectors in the application, this approximation will not have a large impact on the inferences. If prior beliefs are more significant, Appendix A.4 presents some sensible and tractable approaches which were tested and offer a route into this difficult problem.

3.4 Inference for Anomaly Detection

So far i) the inferential paradigm has been discussed, ii) a flexible set of modelling tools to organise beliefs about a single observation have been set out and iii) a method for making inferences about the population to which the observations belong developed. In this section, a novel method for making inferences about the actual quantities of interest in anomaly detection will be presented. For the example in this chapter, that is deciding whether a new observation is anomalous. The quantity of interest is

$$A = \begin{cases} 1 & \text{if the sample is anomalous and} \\ 0 & \text{otherwise} \end{cases}$$

which describes beliefs about the population (anomalous or non-anomalous) to which an entity belongs.

3.4.1 Discrepancy for Second Order Exchangeable Random Fields.

In the examples presented, the expert was not able to specify beliefs about relative population size or an observation given it belonged to the anomalous population, therefore, these cases will not be considered in the main body of the text. For discussion on these points, see Appendix A.4.4. A new test observation is notated by $z^{(i)}$ where $i > N$. Given that it belongs to the non-anomalous population, it is assumed that $z^{(i)}$ is part of the same second order exchangeable sequence of multivariate random fields as the training-set. Then the observation may be written

$$z^{(i)}(t) = h(t)(M(\beta) \oplus R^{(i)}(\beta)) \oplus M(\mathcal{W}(t)) \oplus R^{(i)}(\mathcal{W}(t)) \oplus \epsilon^{(i)}_{me}(t).$$

Two types of anomalies will be considered, global and local.

1. A global anomaly is when the global component is different from expected, given it belongs to the non-anomalous population, which constitutes a low frequency anomaly.

2. A local anomaly is when the residual process is different from expected, given it belongs to the non-anomalous population, which constitutes a high frequency anomaly.

If the interest is in only local or global anomalies, separate tests increase the specificity of the approach. Given the observation is assumed to be part of an infinite set of second order exchangeable random fields, the global and local anomalies can be assessed separately using discrepancy measures for the local and global components. The global discrepancy that would be calculated if all components were known is

$$Dis(R^{(i)}(\beta_{(:)})) = R^{(i)}(\beta_{(:)})^T \left(\mathbb{V}ar[R^{(i)}(\beta_{(:)})]\right)^{-1} R^{(i)}(\beta_{(:)}).$$

As $R^{(i)}(\beta_{(:)})$ and $\mathbb{V}ar[R^{(i)}(\beta_{(:)})]$ are not known, they are replaced with the estimators $\mathbb{E}_{z^{(i)}}[\beta^{(i)}_{(:)}] - \mathbb{E}_{\bar{\beta}}[M(\beta_{(:)})]$ and $\mathbb{E}_{\hat{\Sigma}_\beta}[M(V_\beta)]$. Given that estimators have been used, the uncertainty in the estimators should be taken into account. Therefore, the global discrepancy measure is defined as

$$\begin{aligned} Dis^G\left(z^{(i)}\right) = &\left(\mathbb{E}_{z^{(i)}}[\beta^{(i)}_{(:)}] - \mathbb{E}_{\bar{\beta}}[M(\beta_{(:)})]\right)^T \\ &\left(\mathbb{V}ar_{z^{(i)}}[\beta^{(i)}_{(:)}] + \mathbb{V}ar_{\bar{\beta}}[M(\beta_{(:)})] + \mathbb{E}_{\hat{\Sigma}_\beta}[M(V_\beta)]\right)^{-1} \\ &\left(\mathbb{E}_{z^{(i)}}[\beta^{(i)}_{(:)}] - \mathbb{E}_{\bar{\beta}_{(:)}}[M(\beta_{(:)})]\right). \end{aligned} \tag{3.4.1}$$

Here it has been assumed that $\mathbb{E}_{\hat{\Sigma}_\beta}[M(V_\beta)]$ has been reshaped into a $(rq \times rq)$ matrix. Similarly, the local discrepancy measure is defined as

$$\begin{aligned} Dis^L\left(z^{(i)}\right) = &\left(\mathbb{E}_{z^{(i)}}[\mathcal{W}^{(i)}(t)] - \mathbb{E}_{\bar{\mathcal{W}}(T)}[M(\mathcal{W}(t))]\right)^T \\ &\left(\mathbb{V}ar_{z^{(i)}}[\mathcal{W}^{(i)}(t)] + \mathbb{V}ar_{\bar{\mathcal{W}}(T)}[M(\mathcal{W}(t))] + \mathbb{E}_{\hat{\Sigma}_{\mathcal{W}}(T)}[M(V_{\mathcal{W}}(t))]\right)^{-1} \\ &\left(\mathbb{E}_{z^{(i)}}[\mathcal{W}^{(i)}(t)] - \mathbb{E}_{\bar{\mathcal{W}}(T)}[M(\mathcal{W}(t))]\right) \end{aligned} \tag{3.4.2}$$

where $\mathbb{E}_{\hat{\Sigma}_{\mathcal{W}}(T)}[M(V_{\mathcal{W}}(t))]$ notates the adjusted expectation of the $(q \times q)$ population covariance matrix. These metrics measure the difference between a test observation and the non-anomalous population and large values are indicative of an anomaly.

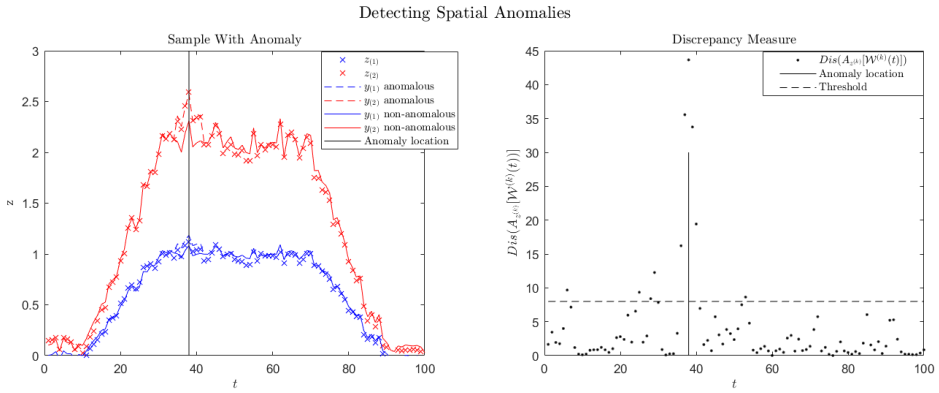

Figure 3.2: Left: sample from the toy model with a local anomaly added. Right: local discrepancy measure for every location.

Figure 3.2 shows a sample from the toy example with a local anomaly which is difficult to spot by eye. To create the anomaly, a sample of $y_{(1)}^{(i)}$ was taken from the toy example and a region T' was set to $y_{(1)}^{(i)}(T') + 0.1$, where $T' = (35, 36, \cdots, 41)^T$. Then $y_{(2)}^{(i)}$ and ϵ_{me} were sampled used equations 2.3.4 and beliefs about the population mean and variance fields were adjusted using 20 samples from the toy example model. Separating the local and global components allows the anomaly to be detected; shown by the large discrepancy values on the right hand plot.

3.4.2 Nested Tests

The expert may have additional beliefs that certain types of anomaly are more likely than others. Furthermore, certain types of anomalies may be less desirable than others. Here two types of anomalies are defined:

1. spatial anomalies where the distribution within a field is different from expected and
2. spectral anomalies where the relationship between constituent fields is different from expected.

It is assumed that one is nested in the other, for example, all spectral anomalies are also spatial anomalies. Two discrepancy metrics are used, one for the spatial anomaly and one for a spectral anomaly. Spatial anomalies are identified and regions which are not spatial anomalies are used to adjust beliefs about the relationship between fields. For example, if the observation is a bi-variate field, the model

$$z_{(2)}(t) = \sum_{m=1}^{d} \alpha_{(m)} h_{(m)}(z_{(1)}(t)) \oplus \epsilon(t)$$

could be used to make the specification. Here α is a $(d \times 1)$ vector of regression parameters and $\epsilon(t)$ residuals. The spectral anomalies are identified using the residual of an observation from their projection onto the line. The line could, for example, be defined by

$$\hat{\epsilon}(t') = z_{(2)}(t') - h(z_{(1)}(t'))\mathbb{E}_{D'}[\alpha]$$

where D' is the set of locations which are identified as not spatial anomalies by the spatial discrepancy measures. This allows the different types of anomalies to be identified. For example, plastic in chocolate would be a spectral anomaly while damaged chocolate would be a spatial one. Figure 3.3 shows a spectral anomaly for the toy example which exhibits a different joint distribution than the non-anomalous population. The red points are spectral anomalies and have drifted off the curve.

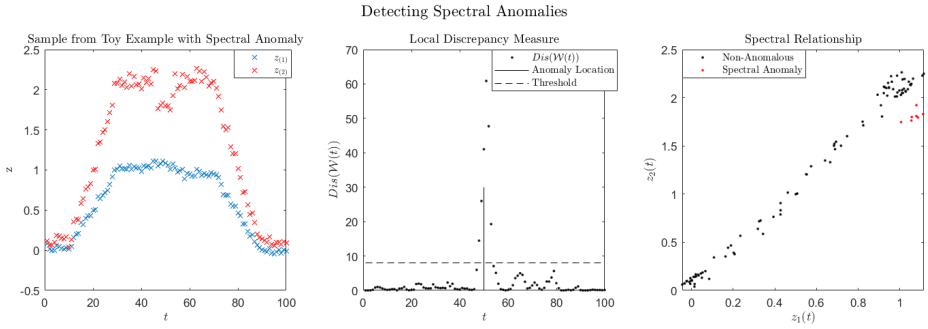

Figure 3.3: Left: sample from the toy example with a spectral anomaly added. Centre: local discrepancy measure for every location in z. Right: $z_{(1)}$ plotted against $z_{(2)}$ with spectral anomaly shown in red.

Before moving onto the industrial application it is worth summarising the key ideas. The aim is to decide whether an entity belongs to the same population as the non-anomalous population. A set of observations of entities known to belong to the non-anomalous population are made. Each field is considered a sample from a population measured imprecisely. Inferences are made about the sample from the population and then about the population to which the samples belong. Then for a test observation, discrepancy metrics are used to decide if their is evidence the test observation belongs to the same population as the non-anomalous training-set. The observations are large and complex. However, at its root, the algorithm is making inferences about a population and deciding whether there is evidence test samples do not belong to that population. This not an unusual idea in statistical practice.

3.5 Finding Plastic in Chocolate

It is now time to use the methodology described in chapters 2 and 3 to address the plastic in chocolate problem, problem statement 1.1.1.

3.5.1 X-ray Preliminaries

Electromagnetic radiation in a certain energy range is referred to as an X-ray. As an X-ray passes through a substance it may interact with the sample in one of three ways:

1. photoelectric absorption, where the incident photon is completely attenuated,
2. Compton (incoherent) scattering, where it changes direction and loses energy, and
3. Rayleigh (coherent) scattering, where it changes direction but the energy does not change.

The rate at which these interactions occur is dependent on the density and chemical composition of the substance traversed by the X-rays. Therefore, if the manner in which X-rays interact with an entity can be deduced, information about the entity's composition can be inferred. X-ray examination is a method of deducing the manner in which X-rays interact with an entity; it measures the total energy of X-rays which have traversed the sample. The fundamental components of an X-ray examination system are i) a source, ii) an entity and iii) a detector. Photons are produced at the source and travel through the entity to the detector which records the energy deposited. When an entity is present, the total energy recorded at the detector is notated I and when it is not I_0. X-ray images of chocolate bars have low scatter-to-primary-ratio which is defined as follows.

Definition 3.5.1. ***Scatter-to-primary-ratio:*** *intensity due to photons that have undergone a scattering interaction divided by the intensity due to photons that have not interacted with the sample.*

Therefore, in this chapter, the scattering interactions may be ignored.

A Mono-Chromatic Model

To explain the modelling decisions made, it is useful to introduce two simple physics models. These are not used directly to make inferences in this chapter but motivated i)

the data processing, ii) the form of the model and iii) the equipment used. A deterministic simulator for a single pixel mono-energetic X-ray system is described by

$$I(E) = I_0(E)e^{-\mu(E)\eta} \tag{3.5.1}$$

where η, referred to as entity thickness, is the length a non-interacting photon must travel while inside the entity to reach the detector, E is the energy of the incident photons and $\mu(E)$ is the mass attenuation coefficient. This coefficient describes the rate at which interactions, between photons and entities, occur and it is unique for a given material [86, 121]. Equation 3.5.1 is a good illustration of a fundamental problem when making inferences about the composition of an entity using X-ray examination. Equation 3.5.1 may to rearranged to give

$$\log(\mathcal{T}(E)) = \log\left(\frac{I(E)}{I_0(E)}\right) = -\mu(E)\eta.$$

Then for any $\mu(E)$ and η, a second thickness and material may be found so that the observed intensities are identical. This puts inferring the $\mu(E)$ and η of an entity in the class of ill-posed problems (see Section 5.1 for further discussion on ill-posed problems). For example, a thin piece of lead can produce an identical intensity as a thick piece of aluminium. Therefore, given a thickness of an alloy of chocolate and plastic, there exists a thickness of chocolate such that the observed intensities are identical. This problem is exacerbated because plastic and chocolate are chemically similar so the range of equivalent chocolate thicknesses are within product variation. This implies that conventional X-ray examination and anomaly detection methodologies cannot find plastic contaminants in chocolate. Figure 3.4 illustrates the ill-posedness of material inference using X-rays. The dots represent simulated (using GEANT4, see Section 4.3.2) intensities for a range of thicknesses and alloys of Polymethyl methacrylate (PMMA) and aluminium (Al). The horizontal line is a simulated observed intensity. Even if there is no measurement error, there are an infinite number of solutions which could have produced this intensity. Therefore, although some uncertainty in thickness may be resolved, the alloy of PMMA and Al may not and hence the problem is ill-posed.

A solution to the ill-posedness of this problem is to make two measurements at different incident energies. Then there are two equations

$$\log(\mathcal{T}(E_1)) = -\mu(E_1)\eta \text{ and}$$
$$\log(\mathcal{T}(E_2)) = -\mu(E_2)\eta,$$

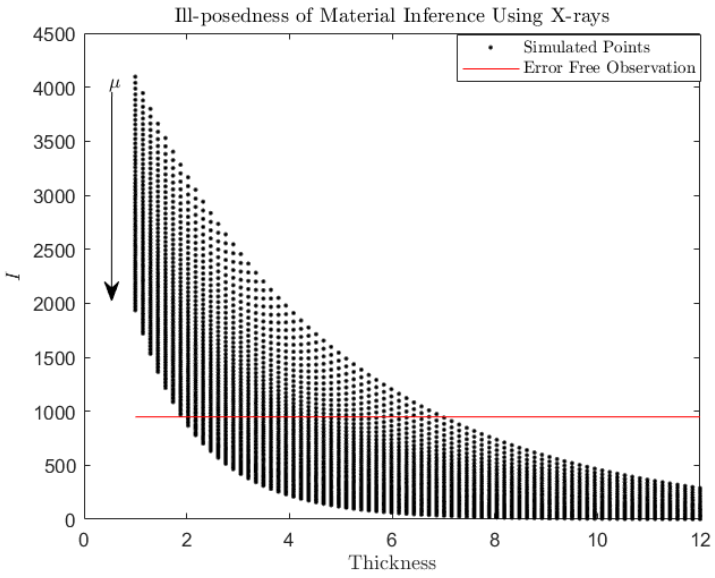

Figure 3.4: Simulated intensities for range of thickness of a range of alloys of PMMA and aluminium. Moving down, the percentage of aluminium in the alloy increases. The red line shows an error free observation to illustrate the ill-posedness of the problem.

and given that the relationship between $\mu(E_1)$ and $\mu(E_2)$ is known for a given material, two unknowns. Consequently, if two measurements are made with two distinct incident energies, the problem is no longer ill-posed.

A Poly-Chromatic Model

Although illustrative, the mono-chromatic model of X-ray absorption is overly simplistic. Examination systems are usually poly-chromatic meaning that the incident X-rays are not a single energy but form a spectrum of energies. To extend the mono-chromatic model to a poly-chromatic version, equation 3.5.2 is adjusted to

$$I = \int_0^\infty N(E) E e^{-\mu(E)\eta} dE \tag{3.5.2}$$

$$I_0 = \int_0^\infty N(E) E dE \tag{3.5.3}$$

where $N(E)$ is the number of photons of energy E and $N(E)E$ is referred to as the incident spectrum. Instead of measurements of two incident energies, measurements at two incident spectra are required to make the problem well-posed. This technique is referred to as multi-spectral X-ray examination and there are two methods for altering the incident spectrum: i) kV switching which alters the upper bound of the spectrum and ii) filtering (beam hardening) which alters the lower end of the spectrum. These techniques

require multiple exposures (measurements) which is time consuming and complications are caused by temporally correlated measurement errors and the alignment of entities changing between measurements.

The systems of interest in this work are Digital Radiography (DR) systems. The detectors are made up of a number of pixels in a square lattice. Then, instead of a single intensity, the observation is a vector of intensities which may be indexed spatially. An X-ray image is then a $(\kappa \times \theta)$ matrix where each element may be indexed by a Cartesian coordinate $t = (i, j)^T$ and the number of locations in an image is $n = (\kappa\theta)$.

3.5.2 MAP Technology

A solution to the problems of using traditional multi-spectral examination techniques is MAP technology, patented by Ibex Innovations [46]. It exchanges using multiple measurements with different incident spectra, taken at two different time points, with varying the incident spectrum spatially. The MAP is a metal sheet with periodically varying thickness which adds a known degree of oscillating beam hardening over a small local area. The change in filtration over a few pixels may be viewed as a change in incident spectrum on the sample. The poly-chromatic model for pixel (i, j) is

$$\begin{aligned} I(i,j) &= \int_0^\infty N(E,i,j)Ee^{-\mu(E,i,j)\eta(i,j)}dE \text{ where} \\ N(E,i,j)E &= N(E)Ee^{-\mu^{MAP}(E)\eta^{MAP}(i,j)} \text{ and} \\ I_0(i,j) &= \int_0^\infty N(E)Ee^{-\mu^{MAP}(E)\eta^{MAP}(i,j)}dE. \end{aligned}$$

Here $I(i, j)$ is the intensity at location (i, j) on the detector, $\mu^{MAP}(E)$ is the MAP attenuation coefficient and $\eta^{MAP}(i, j)$ is the thickness of the MAP at pixel (i, j). Given that the sample in a local region is uniform, a local clique of pixels may be used as a proxy for mutli-spectral observations.

The next challenge is to extract the multi-spectral information present in a single image. The novel method − developed specifically for MAP technology − by which the multi-spectral information is extracted is referred to as continuous energy. Due to engineering complications, it is impracticable to produce a MAP and detector such that each pixel is exposed to a single MAP thickness. Due the discretisation process of DR, the detector under-samples the MAP creating a signal with a continuum of beam-hardening.

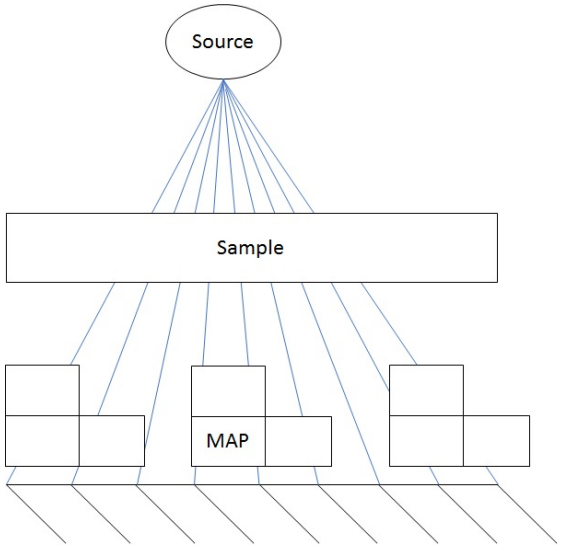

Figure 3.5: A diagram of a MAP equipped X-ray examination system. The blue lines show the effective area of MAP averaged by each pixel. The stacked blocks are the thicknesses of MAP present in the system.

Figure 3.5 depicts a diagram of a one dimensional detector with a three level MAP. A level refers to the thickness of material used to make up the MAP's structure. The area seen by each pixel is shown by the blue lines. Each pixel sees a variable amount of MAP level so it is illogical to classify each pixel into one of the three levels present. The continuous energy method models the relationship between energy incident on the entity and the energy recorded by the detector with

$$\log(\mathcal{T}(t)) = \beta_0(t) + \beta_1(t)I_0(t). \tag{3.5.4}$$

Assuming an error-free observation, for any thickness of any alloy of two known materials, β is unique and the relationship between material and thickness can be written as $\mathcal{F}(\eta, \mu) = \beta$ where $\mathcal{F}$ is a function. Figure 3.6 shows that the problem of composition inference is no longer ill-posed when continuous energy parameters are available. The plots depicts simulated continuous energy parameters for a number of entities with known composition. Continuous energy is the appropriate way of extracting multi-spectral information with a MAP. Furthermore, it has advantages over traditional techniques as it is more invariant to spatial variation across the detector. The estimates for β, notated β^{EC}, are calculated for every pixel on the detector resulting in two images.

Given that it has been shown in simulated models that the problem is no longer ill-posed, the next step is to calculate the parameters in equation 3.5.4 in the real world for every pixel on a detector. Real world images contain measurement error which is closely

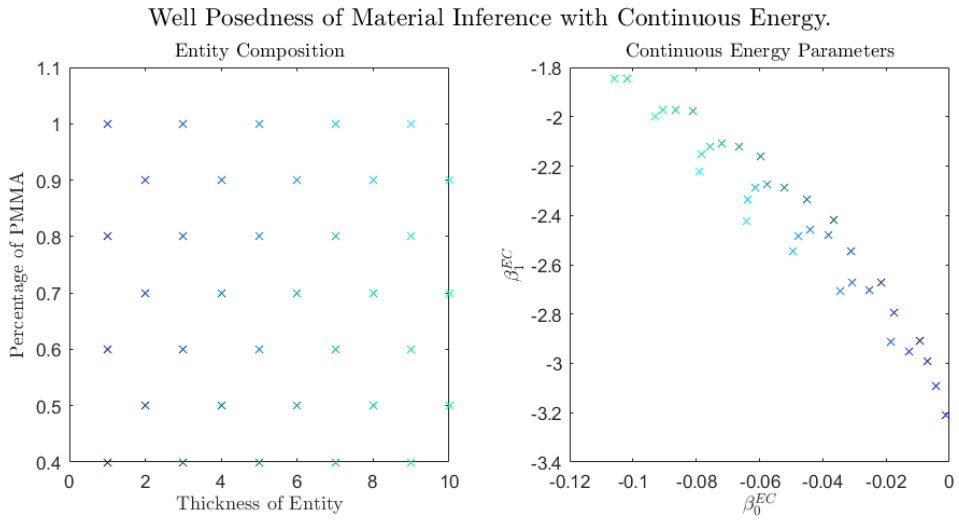

Figure 3.6: Left: (μ, η) for a set of alloys of PMMA and aluminium. Right: simulated continuous energy outputs for the set of alloys and thicknesses on the left.

related to Poisson noise. This is because the photons hitting the detector is a Poisson process which is then averaged by the digital detector creating a integral over Poisson processes [121]. The pseudo GLS estimates, notated $\beta^{EC} = (\beta_0^{EC}, \beta_1^{EC})^T$, are calculated using a local clique of an image without an entity and the same clique of an image with the entity. This is repeated for every pixel on the detector and the result is two images one of estimates $\beta_0^{EC}(t)$ and the other $\beta_1^{EC}(t)$. This can be done efficiently using the fast Fourier transform and the convolution theorem. For details on how the MAP is processed to produce a continuous energy estimate at every pixel, see Appendix A.5.

The Observations

For the chocolate bar example, the continuous energy parameters were calculated at every pixel on the detector. In this section, the observation are taken to be the model dependent continuous energy parameters. The observation at location t is notated

$$z(t) = \beta^{EC}(t)$$

which is a vector of elements on a bi-variate spatial field. The algorithm was trained on a set of observations and then tested on a second set. The training-set was twenty observations of chocolate bars known to be from the non-anomalous population. To test the algorithm a further set of forty chocolate bars were used. Seven bars were contaminated with one or more pieces of polycarbonate or polystyrene. The contaminants were inserted in the bars using a soldering iron. Control regions were also added using the soldering iron

to melt chocolate where no plastic was added. This was to ensure the signal being produced was due to plastic and not to the change in composition caused by burnt chocolate.

The Model

The model for the chocolate bar was organised as the independent sum of two random fields, that is observation with index i is written

$$z^{(i)} = y^{(i)} \oplus \epsilon_{me}^{(i)}$$

where the measurement error was assumed to have a quadratic shape with an additional independent noise term i.e.

$$\begin{aligned}
\epsilon_{me}^{(i)} &= \epsilon_{me}^{G(i)} \oplus \epsilon_{me}^{L(i)} \text{ where} \\
\epsilon_{me}^{G(i)}{}_{(k)}(t) &= (1, t_{(1)}, t_{(2)}, t_{(1)}^2, t_{(2)}^2)\beta_{(:,k)}^{me(i)} \\
\mathbb{V}ar[\epsilon_{me}^{L(i)}(t)] &= \Sigma_{me}^{(i)} \text{ and } \mathbb{C}ov[\epsilon_{me}^{L(i)}(t), \epsilon_{me}^{L(i)}(t')] = 0.
\end{aligned}$$

This measurement error model was chosen because the number of photons produced by the source varies spatially and temporally. A quadratic shape accounts for the change in the number of photons spatially and a different quadratic would be produced in each observation. The random error component is to model the Poisson noise in the image which is assumed to be uncorrelated spatially. As calculating z uses a moving window, this is not strictly true but the simplification does not have a large impact on the inferences performed. The measurement error model was validated by checking the mean of the $z^{(i)} - H\hat{\beta}^{me(i)}$ looked uncorrelated so that the exchangeability assumptions were reasonable. In other words, any systematic temporal drift had been removed. The set $\{y^{(1)}, \cdots, y^{(N)}\}$ are assumed to be part of a second order exchangeable sequence of multivariate random fields. This implies that

$$\begin{aligned}
y^{(i)}(t) &= \sum_{\omega=1}^{\tilde{q}} \pi_{(\omega)}(t) y_{(:,\omega)}^{(i)}(t) \text{ where} \\
y_{(:,\omega)}^{(i)}(t) &= M^G(y_{(:,\omega)}(t)) \oplus M^L(y_{(:,\omega)}(t)) \oplus R^{G(i)}(y_{(:,\omega)}(t)) \oplus R^{L(i)}(y_{(:,\omega)}(t)) \\
&= \mathcal{M}_{(:,\omega)}^{(i)}(t) \oplus \mathcal{W}_{(:,\omega)}^{(i)}(t).
\end{aligned}$$

Here

$$\begin{aligned}
\mathcal{M}_{(k,\omega)}^{(i)}(t) &= M^G(y_{(k,\omega)}(t)) \oplus R^{G(i)}(y_{(k,\omega)}(t)) \text{ where} \\
M^G(y_{(k,\omega)}(t)) &= (1, t_{(1)}, t_{(2)}) M(\beta_{(:,k,\omega)}) \text{ and} \\
R^{G(i)}(y_{(k,\omega)}(t)) &= (1, t_{(1)}, t_{(2)}) R^{(i)}(\beta_{(:,k,\omega)}).
\end{aligned}$$

The chocolate bar in question is a rectangle with two truncated square pyramids sticking out. Therefore, the target shape of the production process would produce an X-ray image with a linear surface for every segment. The production process is not perfect and the X-ray examination procedure involves a projection onto a plane. Therefore, a residual process was required to incorporate the behaviour which is difficult to model with a linear combination of basis functions. The local component of the random field is given by

$$\mathcal{W}^{(i)}_{(k,\omega)}(t) = M^L(y_{(k,\omega)}(t)) \oplus R^{L(i)}(y_{(k,\omega)}(t)).$$

The constituent fields of the mixture were assumed to be independent. Each constituent field had a correlation function and partition specific variance hyper-parameter, notated $\sigma^{(i)}_{(k,\omega)}$. The correlation function is given by equation 2.2.8 with $p = 2$. Enforcing continuity over boundaries would be a better representation of beliefs about the object. However, there was a large computational cost and, given that the inferential procedure must not slow down the manufacturing process, the additional cost was not deemed worthwhile. This model was not chosen in isolation but by validating the model on individual chocolate bars. The model was fitted on a number of bars − which were from the non-anomalous population or contained plastic − analysing diagnostic measures (see Section 2.2.2), examining the adjusted expectations and checking that contamination was more visible after the expectation of the global component was removed.

The Training Procedure

The regression parameters and variance hyper-parameters were fixed to their GLS estimates and the correlation lengths were fixed to $l = 10$. GLS estimates were chosen in this example because there was a large amount of data available and prior beliefs were having little impact on adjusted beliefs. The experiment was considered to be well enough controlled that a single set of control points was sufficient for all observations and uncertain partitions were not necessary for this example. As equipment used in factories is unlikely to be as well controlled as this, this work should be viewed as a proof of concept as, if the problem cannot be solved when it is reasonable to assume known partitions, a solution with uncertain partitions is unlikely to be forthcoming. Figure 3.7 left shows the segmentation used and right shows an image of $z^{(1)}_{(1)}$. The segmentation was produced in three stages. First, use a threshold to classify the region without chocolate. Second, specify the vertexes (red crosses in Figure 3.7) and the edges between them; any point

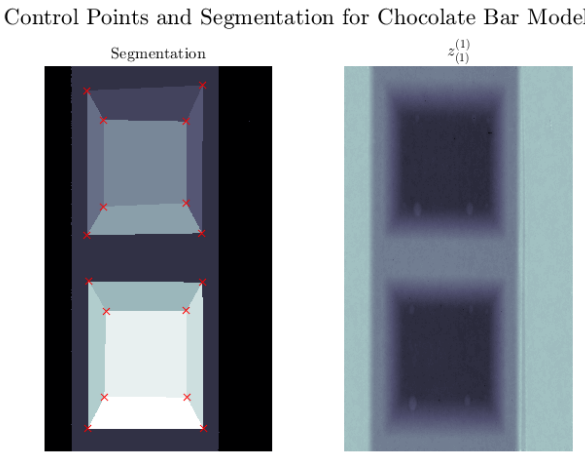

Figure 3.7: Left: segmentation of observations for mixture of multivariate random fields model. The red crosses indicate the control points. Right: $z_{(1)}^{(1)}$ image for the first entity in the training-set.

within a closed loop is within a segment. Finally, any region which is chocolate but is not within a closed loop is in the segment describing the rectangular base.

Observation Learning

The belief adjustment for $\mathcal{W}_{(k,\omega)}^{(i)}$ was performed in two stages: the regression component was adjusted using 1000 points from the field and the residual process was adjusted using approximate kernel methods. The approximate kernel for the residual process update was designed using equation 2.4.3 with a window size of 30. The population variance and correlation fields were adjusted separately. The covariance specification was made using the correlation function, given in equation 2.2.8, with $p = 2$ and $l = [10, 10]$. Both population variance and correlation fields were assumed to be stationary, therefore the two stage approach was not required.

A region of the image, black in Figure 3.7, is referred to as open beam. This is the proportion of the detector that has no chocolate bar in front of it. Therefore, this area has no uncertainty induced by the variation in the chocolate bar population and it may be used to adjust beliefs about the measurement error. The population mean of this region is known to be $[0, 0]$.

In summary, the processing algorithm for every observation is as follows.

1. Using the open beam, black in Figure 3.7 left, calculate the GLS for $\beta^{me(i)}$ (key formula 2.2.1) and Σ_{me}.

2. Take the expected measurement error away from the observation using

$$z^{(i)} - H\hat{\beta}^{me(i)}$$

where $\hat{\beta}^{me(i)}$ is the GLS estimator for $\beta^{me(i)}$.

3. Then for partition ω and constituent field k.

   (a) Take a sample of 1000 pixels from the partition and calculate the GLS estimates for $\beta^{(i)}_{(:,k,\omega)}$ and variance multiplier $\sigma^{(i)}_{(k,\omega)}$ (key formula 2.2.1).

   (b) Assuming the adjusted expectation of $\beta^{(i)}_{(:,k,\omega)}$ is given by the GLS estimate, take away the expected global variation component and produce an approximate kernel (equation 2.4.3).

   (c) Calculate the adjusted beliefs about the local component of the field using the approximate kernel.

   repeat for every partition ω (except the open beam) for both $z^{(i)}_{(1)}$ and $z^{(i)}_{(2)}$.

Population Learning

The aim of this study was to assess whether it is possible to detect small plastic contaminants and, for this reason, global anomalies were not assessed in this example. The global component still has an important role as it removes systematic variation making it easier to assess whether there is anomalous local variation. There are many components of the algorithm presented here which are tuned to detecting plastic contaminants. If this algorithm is used in industry, analysts can add further components to detect other types of anomalies. For example, a global anomaly could constitute a bar which does not meet other quality control criteria. We will show the methodology presented is able to separate out these different types of anomaly and each type may be assessed separately or together.

To adjust beliefs about the population mean and variance of the local component, the sample mean and the sample covariance of the adjusted expectations of the local component were calculated at every location. This gives five images, two informative for the population mean, two informative for the population variance and one informative for the population correlation at every location. Each image is adjusted separately using an approximate kernel calculated using equation 2.4.1 with a window size of 30. The specification of the population mean field is calculated using equation 3.3.2 and the population

variance components using equations 3.3.3 and GLS approximations given in equation 3.3.4.

In summary, the population learning procedure is as follows.

1. Calculate the sample mean vector and covariance matrix for every location.
2. Reorganise the elements into five images.
3. Then for all five images
   (a) calculate the approximate kernels and
   (b) calculate the approximate adjusted expectations.

Figure 3.8 illustrates the population learning procedure. The top line shows the sample mean, variance and correlation estimators. The middle line shows the approximate adjusted beliefs. The bottom line shows some of the approximate kernels used in the updating procedure. Table 3.1 summarises all the specifications which were made in order to perform the analysis, here $T_{1(\omega)}$ and $T_{2(\omega)}$ are the sample mean of the residual process GLS approximated adjusted variances.

Table 3.1: Table of specifications for finding plastic in chocolate.

Quantity	$\mathbb{E}$	$\mathbb{V}ar$	p	l
β^{me}	GLS	GLS	N/A	N/A
Σ^{me}	GLS	N/A	N/A	N/A
$\beta^{(i)}_{(:,k,\omega)}$	GLS	GLS	N/A	N/A
$R^{(i)}(\mathcal{W}_{(k,\omega)}(t))$	0	$\sigma^{(i)}_{(k,\omega)} - \mathbb{V}ar[M(\mathcal{W}_{(k,\omega)}(t))]$	2	[10,10]
$\sigma^{(k)}_{(k,\omega)}$	GLS	N/A	N/A	N/A
$M(\mathcal{W}_{(1,\omega)}(t))$	0	1×10^{-6}	2	[10,10]
$M(\mathcal{W}_{(2,\omega)}(t))$	0	0.002	2	[10,10]
$M(V_{\mathcal{W}(1,\omega)}(t))$	4×10^{-6}	3×10^{-10}	2	[10,10]
$M(V_{\mathcal{W}(2,\omega)}(t))$	0.001	0.001	2	[10,10]
$M(C_{(1,2,\omega)}(t))$	0	0.06	2	[10,10]
$\hat{\Sigma}_{\mathcal{W}(1,\omega)}(t)$	$4 \times 10^{-6} + T_{1(\omega)}$	$3 \times 10^{-10} + 1 \times 10^{-11}$	N/A	N/A
$\hat{\Sigma}_{\mathcal{W}(2,\omega)}(t)$	$0.001 + T_{2(\omega)}$	$0.001 + 1 \times 10^{-5}$	N/A	N/A
$\hat{C}_{(1,2,\omega)}(t)$	0	$0.06 + 1 \times 10^{-5}$	N/A	N/A

Testing a New Observation

To detect plastic contaminants, a two stage nested test was applied. The first test identified high frequency spatial errors using the residual process discrepancy measure in equation 3.4.1. Any pixel with local discrepancy greater than 26.6 was considered a spatial anomaly. This cut off was specified using the sample mean and variance of the local discrepancy of a single observation from the training-set and the three sigma rule [47]. The local discrepancy image was then segmented into connected anomalous regions. A map of the sign of the error is computed using $sign(\mathbb{E}_{z_{(1)}^{(i)}}[\mathcal{W}_{(1)}(T)] - \mathbb{E}_{\bar{\mathcal{W}}_{(1)}}[M(\mathcal{W}_{(1)}(T))])$. Any connected group with over two thousand connected components of the same sign is considered a spatially anomalous region and the second test is applied to that region. Contamination causes a region to drift from the population mean in the same direction, therefore small regions that have the same sign error are grouped together. This step prevents two connected regions of opposite sign from averaging each other out in the second test.

The second test is calculated using the residual from the line

$$z_{(2)}(t) = \left(1, z_{(1)}(t), z_{(1)}(t)^2\right) \alpha \oplus \epsilon \tag{3.5.5}$$

where the residuals ϵ were assumed to be uncorrelated spatially, have expectation zero and constant variance. Given the entity was made of 100% chocolate, ideally the relationship inferred would be $\mathcal{F}(\eta(t)) = z(t)$ where $\eta(t)$ is the thickness of chocolate at location t. Then, if $z(t)$ is measured without error and $\mathcal{F}$ is known, $\eta(t)$ could be identified precisely from either $z_{(1)}(t)$ or $z_{(2)}(t)$. As η is unknown, the relationship is challenging to infer directly, however, this argument implies that $(z_{(1)}(t), z_{(2)}(t))$ should sit on a curve. Therefore, by making inferences about α in equation 3.5.5, that curve can be inferred and, if there is evidence that $z(t')$ does not sit on that curve, it implies that $z(t')$ is not 100% chocolate. Given there is measurement error and a contaminant would cause a group of adjacent pixels all to drift off the curve, regions identified as a connected spatial anomaly are tested together. This enables the effect of measurement error to be reduced and sensitivity to be increased.

Once spatial anomalies were identified, the GLS estimates for α, notated $\hat{\alpha}$, were calculated for each spatially anomalous region using a surrounding clique of pixels which are identified as not spatial anomalies. The spectral test was then whether the distribution of residuals from the inferred line, in equation 3.5.5, had a median greater than zero. As both constituents of $z(t)$ are polluted with noise, instead of taking the vertical residual, a projection onto the line in the direction of least variance was taken. The residual is

defined as

$$\hat{\epsilon}(t^{(k)}) = \text{sign}\left(z_{(1)}(t^{(k)}) - h(z_{(2)}(t^{(k)}))\hat{\alpha}\right)$$
$$\min_{l}\left\{\left(z(t^{(k)}) - (z_{(1)}(t^{(l)}), h(z_{(2)}(t^{(l)}))\hat{\alpha})^T\right)^T\left(\hat{\Sigma}_{me}\right)^{-1}\left(z(t^{(k)}) - (z_{(1)}(t^{(l)}), h(z_{(2)}(t^{(l)}))\hat{\alpha})^T\right)\right\} \tag{3.5.6}$$

where $l = (1, \cdots, \tilde{N})$ is the set of locations in the local clique used to calculate the GLS estimate for α. Then the metric used to decide if plastic contamination was present was

$$\mathbb{I}_{\text{median}\left(\hat{\epsilon}(t^{(1)}),\cdots,\hat{\epsilon}(t^{(N')})\right)>0}\frac{\text{median}\left(\hat{\epsilon}(t^{(1)}), \cdots, \hat{\epsilon}(t^{(N')})\right)}{\sqrt{\frac{1}{N'}\left(\frac{1}{N'-1}\sum_{k=1}^{N'}\left(\hat{\epsilon}(t^{(k)}) - \frac{1}{N'}\sum_{l=1}^{N'}\hat{\epsilon}(t^{(l)})\right)^2\right)}} \tag{3.5.7}$$

where N' is the number of pixels in a connected region of spatial anomalies. The logic behind this measure is that all locations which have pure chocolate should sit on the line defined by equation 3.5.5. If a region also contains another type of material, the points in that region will shift onto a different line. Therefore, if the centre of this distribution of $\hat{\epsilon}(t^{(k)})$ for a region is different from zero, it is evidence that contamination is present. Both types of plastics considered are less dense than chocolate. Therefore, plastic contamination would induce a positive residual $\hat{\epsilon}(t^{(k)})$ and, if plastic contamination is present, the centre of mass of the distribution would be positive. If other types of contamination are present, they may also be assessed. However, if only plastic contamination is of interest, only considering positive residuals will increase the specificity of the approach. The median was chosen over the mean as it provided better separation between contaminants and non-contaminated spatial anomalies. We believe this was because the affect of outliers was reduced.

In summary, the testing observation is processed in the same way as a training observation and the testing procedure is as follows.

1. The local discrepancy measure, given by equation 3.4.2, is calculated for every pixel.

2. Any locations with local discrepancy over 26.7 are identified as spatial anomalies.

3. Sets of connected spatial anomalies are grouped together into a region of spatial anomalies. Any region with under two thousand connected components is discounted as too small to be of interest.

4. For all regions with over two thousand connected components:

(a) the GLS estimate for α (in equation 3.5.5) is calculated using a surrounding clique of pixels,

(b) for every pixel in the region, the residual to that curve (equation 3.5.6) is found and

(c) the spectral measure (equation 3.5.7) for that region calculated.

In order to ascertain the distribution of 3.5.7 for non-contaminated spatial anomalies, a sampling approach was adopted. The approach is as follows: i) randomly sample an observation from the training-set, ii) sample a location on that observation, iii) sample a radius. Then all points within the radius of the sampled location are considered a spatial anomaly. The testing algorithm is then run on that spatial anomaly and 3.5.7 calculated. This was repeated 1000 times and the 95% and 99% quantiles are estimated from those 1000 samples of the measure. Ideally only regions which are identified as spatial anomalies would be sampled, however, there were not enough spatial anomalies in the training-set to produce an estimate of the quantiles so this approach was considered a reasonable compromise.

The final stage of the training procedure was cross validation. Each training observation is set aside and used as a test observation. Then beliefs are adjusted using the remaining 19 observations and the testing procedure applied to the observation that has been set aside. This tests the full algorithm on bars known to be from the non-anomalous population. Figure 3.9 shows the spectral discrepancy measure for every region identified as a spatial anomaly in the cross validation procedure.

Figure 3.10 shows a region identified as a spatial anomaly which is not identified as a spectral anomaly. Figure 3.11 shows a region identified as a spectral anomaly. The difference between the spatial anomaly and spectral anomaly is illustrated best by the histograms. The centre of mass of the spectral anomaly has drifted significantly further from zero indicating correctly that the spatial anomaly is caused by contamination.

3.5.3 Results

Figure 3.12 shows the spectral discrepancy measure for every region in the test set that was considered a spatial anomaly. The ten largest values were the contaminated regions. All contaminants were identifiable at the 95% threshold and six were identifiable at the 99% threshold. No spatial anomalies were incorrectly identified as contaminated. All

contaminants and control regions were identified as spatial anomalies. Figure 3.13 shows images of $z^{(i)}_{(2)}$, local discrepancy and spectral discrepancy measures for the contaminated bars. The colours in the final row indicate the spectral discrepancy with white being over the 95% decision threshold. Test observation four has a control and a plastic contaminant. The control is a more significant spatial anomaly but the spectral test is able to identify the difference between controls and contamination.

3.6 Conclusions

In this chapter, the aim was to use the methodology for modelling multivariate random fields to answer problem statement 1.1.1; detect small plastic contaminates in chocolate bars. The aim of the study was to assess whether MAP technology could be used to provide an automated solution to this problem. To our knowledge, this is the first time X-rays have been able to detect plastic contaminants above product variation and the results are seen as a successful proof of concept. MAP technology has a pivotal role; it enables the difference between a change in thickness of chocolate and the addition of plastic to be distinguished. The method essentially has two stages: spatial modelling to identify suspicious regions and a test to identify if a suspicious region contains a contaminant. The spatial modelling step is stable in that regions with contamination are always clearly identified as a spatial anomaly. The second test is key as it allows the differentiation of bubbles and contaminants which increases the specificity of the approach.

The spatial modelling component of this algorithm has a number of steps which have a logical interpretation. For each observation in the training-set, low frequency behaviour is inferred and then removed to produce an image which is approximately flat. Then a smoothing process is performed to reduce the effect of measurement error. The sample mean and covariance is calculated at every location producing five images. Another smoothing process is performed on the five images to reduce the effect of estimator error. Then a distance between expected and revealed is calculated for a test sample. These ideas are common in many fields, for example signal processing. Bayesian methods are often seen as inaccessible and arcane by non-statistical experts so showing how an expert's beliefs transform into operations on data they are familiar with is useful for building confidence in the methods and facilitating their use.

As plastic and chocolate are similar materials, the signal from contamination is small. For a viable solution in industry, the number of bars which can be incorrectly identified as

anomalous was reported to be one in ten thousand. There were thirty-one spatial anomalies identified − which were not caused by a control or plastic contamination − in the cross validation and test sets. Therefore, the rate at which spatial anomalies occur is around once every two bars. If one in ten thousand false positive identifications is required, the rate at which non-anomalous bars are identified as anomalous would have to be increased. There are two ways this may be done: decrease the rate at which spatial anomalies not caused by contamination are identified or increase the signal from contamination increasing the specificity of the spectral test. The equipment used for this study was high resolution laboratory X-ray equipment. The equipment traditionally used in factories is lower resolution with higher noise and spatial artefacts on the image. Furthermore, the bars will not be as well aligned as in the experimental data presented. Therefore, there are still challenges before the method can be transferred to a complete industrial solution. The method is in the process of being adapted for use on industry relevant equipment and the hope is to incorporate it into an industrial product in the next years.

The next chapter will move onto using multivariate random fields and Bayes linear methods to adjust beliefs about a very different construct, a complex computer model. The aim is to provide a solution to problem statement 1.1.2; given the chemical composition of an entity being examined, infer the effect of scatter photons on an X-ray image. Although the things about which inferences are made are very different, the methods and formalism are similar exemplifying the flexibility of the Bayes linear approach to uncertainty quantification of multivariate random fields.

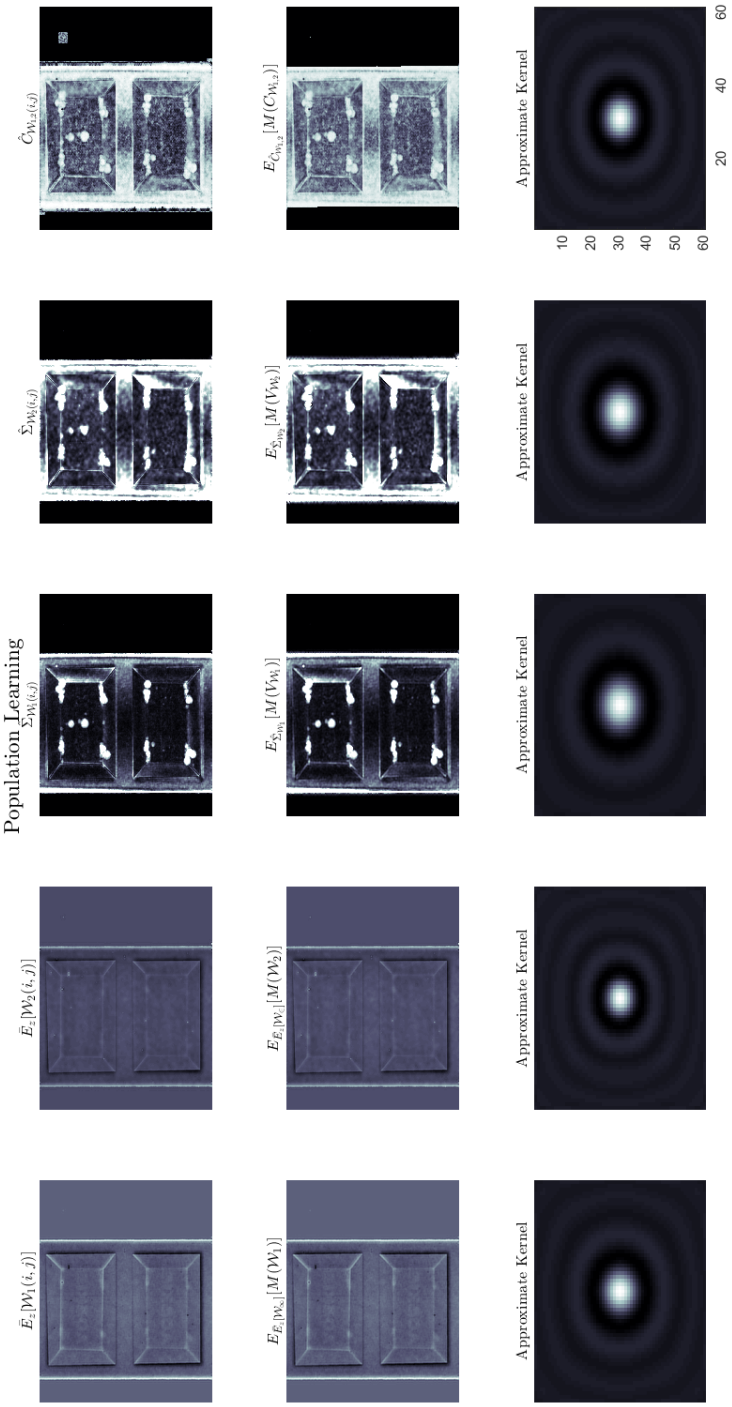

Figure 3.8: Left: unbiased estimators for population mean, variance and correlation. Middle: adjusted beliefs about population mean, variance and correlation fields. Right: sample approximate kernels for the updates above.

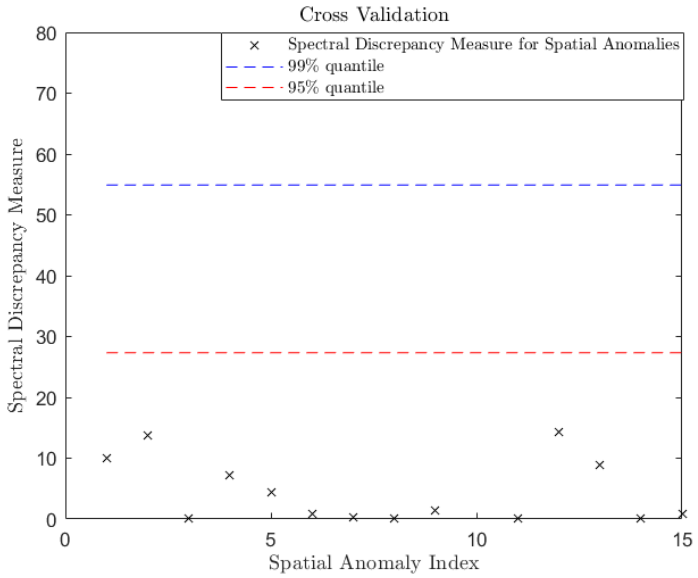

Figure 3.9: Spectral discrepancy measure for spatial anomalies identified in cross validation.

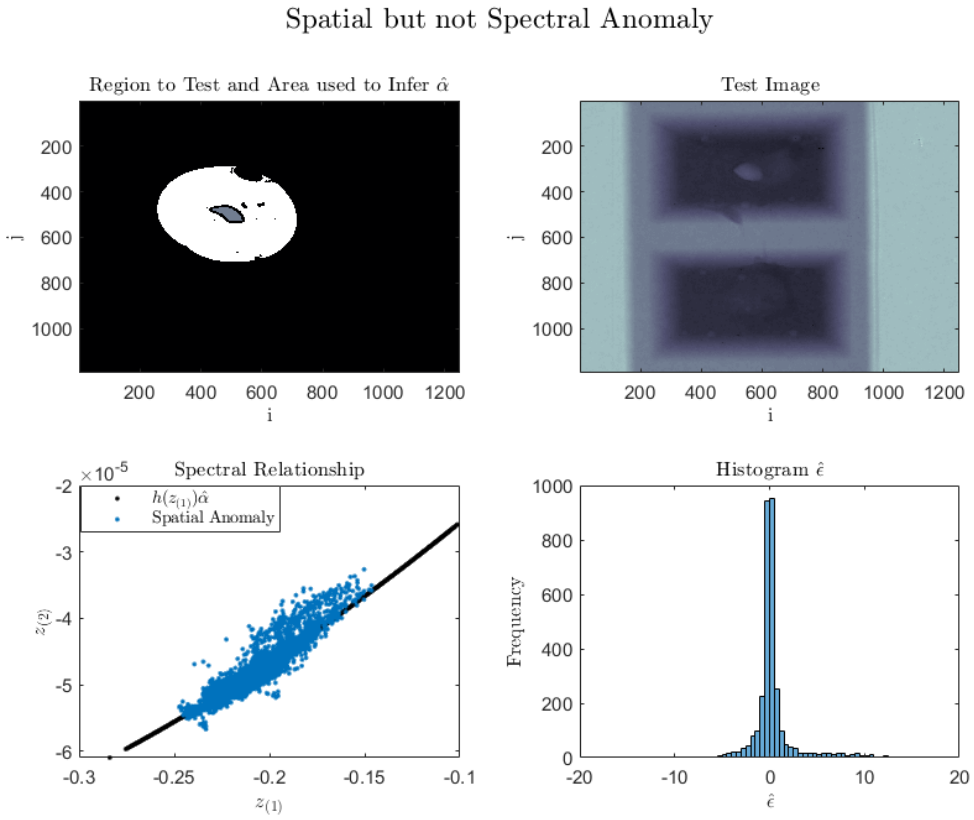

Figure 3.10: Top left: image of spatial anomaly (grey) and area used to infer α (white). Top right: image that is being tested. Bottom left: inferred relationship between $z_{(1)}$ and $z_{(2)}$, and points in spatial anomaly shown top left. Bottom right: histogram of all $\hat{\epsilon}$s in the spatially anomalous region.

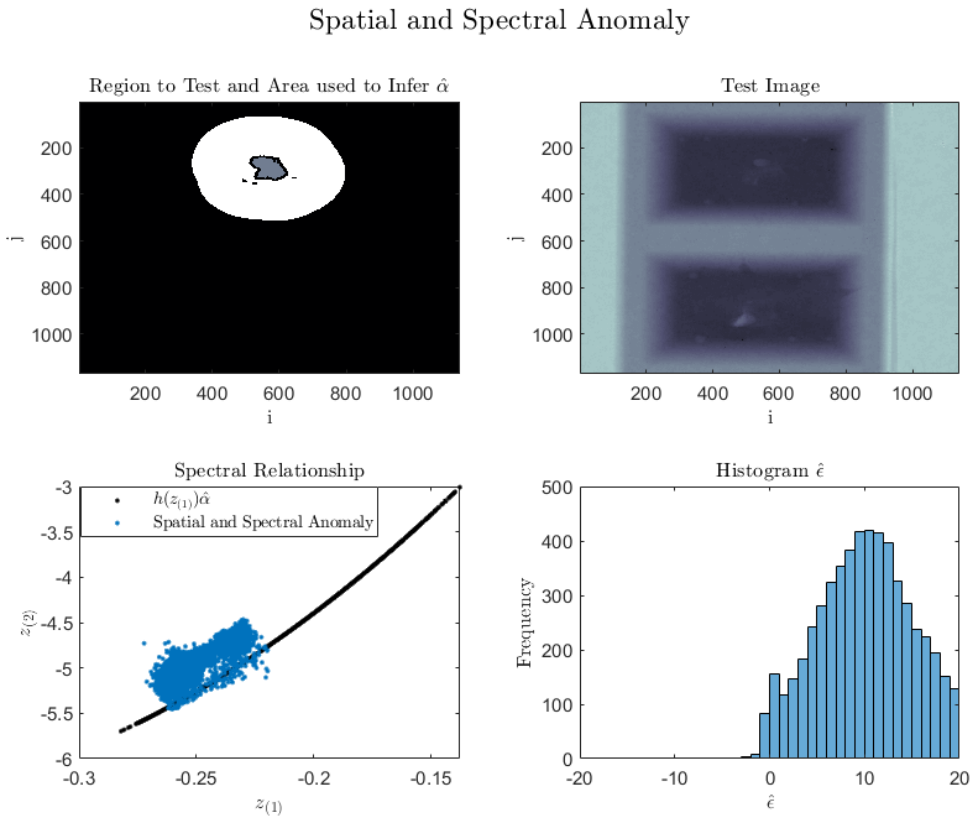

Figure 3.11: Top left: image of spatial anomaly (grey) and area used to infer α (white). Top right: image that is being tested. Bottom left: inferred relationship between $z_{(1)}$ and $z_{(2)}$, and points in spatial anomaly shown top left. Bottom right: histogram of all $\hat{\epsilon}$s in the spatially anomalous region.

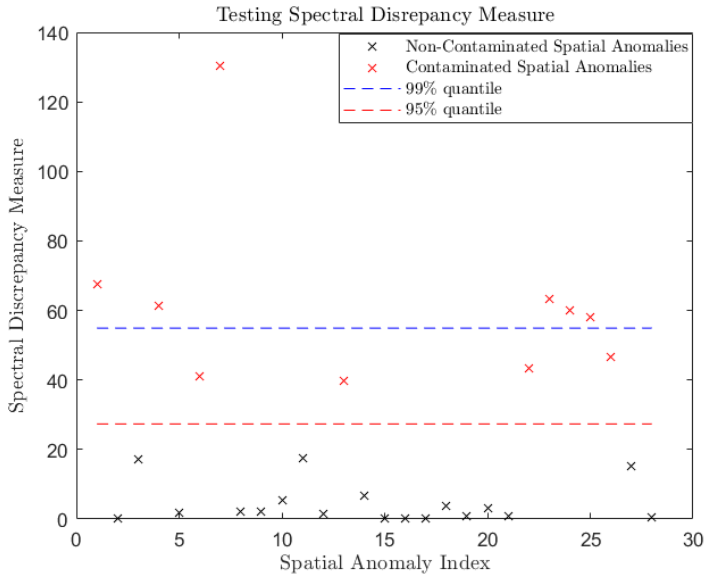

Figure 3.12: Spectral discrepancy measure for all areas identified as a spatial anomaly on every bar in the test set.

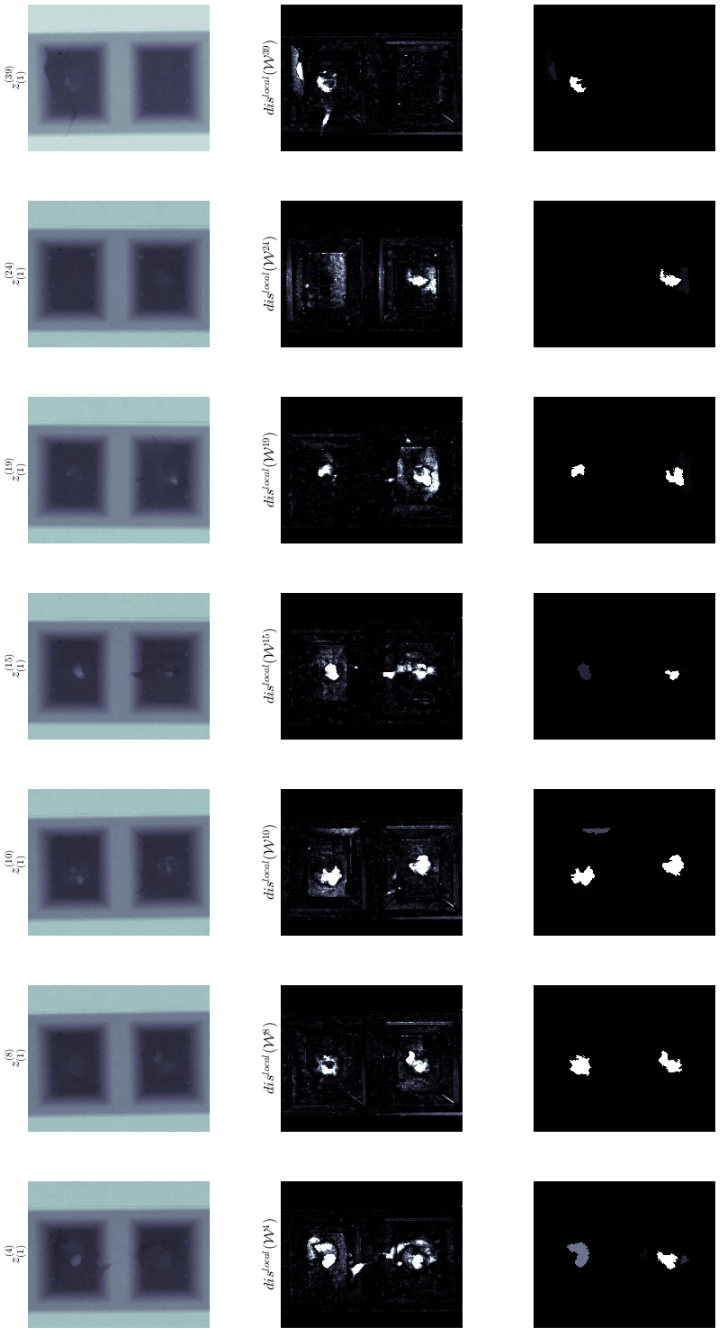

Figure 3.13: First column: $z_{(1)}^{(k)}$ for all images which contain contamination. Second column: images displaying the pixel wise discrepancy of the residual process. Bottom column: areas identified as a spatial anomaly where the white indicates it is also a spectral anomaly. The shade of grey indicates how close to the discrepancy threshold a region is.

Chapter 4

Bayesian Emulation of Imaging Systems

A problem in medical X-ray imaging is caused by photons which have undergone a scattering interaction. They degrade image quality sufficiently that the ability of clinicians to make a diagnosis is impeded. There is an engineering solution called an anti-scatter-grid, but it has two major drawbacks: some scattered photons make it through the grid and, in order to maintain diagnostic quality signal-to-noise-ratio, the dose to the patient has to be increased. As such, a method to accurately predict the intensity present in an image due to scattered photons would have a high impact on medical practice.

There have already been a number of attempts at software based scatter correction. We will separate these attempts into two categories: image processing and physics models. Image processing methods, for example [72], essentially remove low frequency components from the image. The problem is that although scatter is usually a low frequency signal, the signal from photons which have not undergone a scattering interaction may also contain low frequency components. These may also be removed which can have a negative impact on image quality. When we attempted this technique, we found that although it can perform well, the consistency of image quality was far from assured. Other methods which use image processing (but no physics model of scatter) [63, 73–76], will fall into a similar trap as there is not enough in the model to ensure the component removed is scatter. Physics model based approaches use Monte Carlo (MC) simulators [62, 64, 66, 67, 69, 71, 77, 78] to be more selective about the components removed from the image. They use the simulator to estimate a scatter point-spread-function (PSF) which is then used to estimate scatter and remove it from the image. The problem is that, to overcome

the computational expense of MC simulation and convolution, they assume the PSFs are stationary or peice-wise stationary and/or isotropic. These assumptions induce an error as, for complex entities, scatter is non-stationary and non-isotropic. Furthermore, the PSFs depend on the material and thickness of the entity being examined. The methods reviewed select their PSFs from a limited set when there should be a continuum of thicknesses and alloys to select from. In this chapter, an improvement to current methods of scatter estimation will be proposed. Bayesian emulation will be used to reduce the computational impediment and quantify the error between predicted and actual scatter. The method will include non-stationary and non-isotropic PSFs which depend on the material and thickness of the entity being examined. It will be shown that our method can estimate scatter more accurately in a useful time-frame.

The method will be based on the superposition principle, which states, the behaviour of a physical system resulting from two or more stimuli may be decomposed into the sum of the response to each individual stimulus. The principle has been used in mathematical modelling of many phenomena: seismology, optics, oil reservoirs and climate. Given that the superposition principle is useful for modelling phenomena, it is logical to expect that it will also be effective when modelling simulators of that system. This principle greatly simplifies the analysis but it is a modelling assumption and will only be approximately true. A crucial consideration is therefore how to minimise the uncertainty induced by the inadequacy and how to quantify its irreducible portion.

This chapter will be organised as follows: Section 4.1 will describe simulators and their uses in scientific research, in Section 4.2, emulators will be discussed and, in the final section, the application of emulation to predicting scatter will be presented.

4.1 Simulators

Computer models (also known as physics models, simulators or forward operators [31]) are powerful tools for learning about physical systems. They enable inferences to be made about processes which are impractical to observe directly. For example, in a simulator we may observe i) the response of an imaging system to a point source [113], ii) a galaxy coalescing [32] or iii) the climate evolve [23]. Simulators have proved invaluable in many areas of scientific research and there are a wide range of methodologies available for their analysis.

In this thesis, a simulator, notated $\mathcal{F}$, is defined as the computer program provided by

expert to analyst. This program typically encodes a physics model in the form of a system differential equations and a solver which is either deterministic or stochastic. There are usually numbers within the code that have to be specified before it will execute; these are defined as inputs. An evaluation of the simulator at input x is given by

$$\mathcal{F}(x)$$

and $\mathcal{F}$ describes the simulator output at all x. Each simulator evaluation may have multiple outputs which are notated by a subscript. Therefore,

$$\begin{aligned}\mathcal{F}_{(t)}(x) &= \mathcal{F}_{(\text{location } t)}(\text{input } x)\\ \text{and } \mathcal{F}(x) &= \left(\mathcal{F}_{(t^{(1)})}(x), \cdots, \mathcal{F}_{(t^{(n)})}(x)\right)^T.\end{aligned}$$

The inputs, or a subset of the inputs, may also be organised as a multivariate random field with respect to the same space as the output. Therefore, $x = \left(x(t^{(1)}), \cdots, x(t^{(\tilde{n})})\right)^T$ where $x(t^{(i)}) = \left(x_{(1)}(t^{(i)}), \cdots, x_{(p)}(t^{(i)})\right)^T$. For example, the input describes an entity to be measured using X-ray examination equipment and the output an X-ray image.

Although the inputs and outputs have been defined in a general setting, the inputs are often related to system values and the outputs are related to features of system behaviour. Therefore, a simulator is often viewed as a map from system values to system behaviour. For example, a simulator could be a map between carbon dioxide levels in the atmosphere and Earth's temperature. When the relationship between simulator and system has been specified, physics models are useful in research for many purposes: scientific investigation, inverse problem solving, calibrated prediction, uncertainty analysis, sensitivity analysis, decision support and experimental design.

1. Scientific investigation is the process of finding system values which cause behaviour that is scientifically interesting. For example, if $\mathcal{F}$ describes thermohaline circulation (THC) in the ocean, the behaviour of interest could be the collapse of THC [23, 41].

2. Inverse problem solving is finding system values which caused past system behaviour of which an observation has been made. For example, given an observation of Earth's temperature over the last 50 years, what were the carbon dioxide levels in the atmosphere.

3. Calibrated prediction is when observations of historical system behaviour are used to reduce uncertainty in system values and therefore uncertainty in system behaviour.

For example, given observations of the Earth's temperature over the last 50 years, what is the expected sea level over the next 50 years.

4. Uncertainty analysis is how uncertainty, for example in system values, propagates through to system behaviour. For example, given that the future levels of carbon dioxide emissions are uncertain, what is the impact on uncertainty about future climate behaviour.

5. Sensitivity analysis is the investigation of how responsive system behaviour is to changes in system values. For example, how future sea levels respond to a change in carbon dioxide production.

6. Decision support is to aid policy which may affect the system. Some of the inputs, referred to as decision inputs, are associated with real world actions. The decision inputs are chosen to optimise favourable behaviour. For example, if a decision input relates to carbon emissions, what decision would reduce carbon emissions enough to prevent THC collapsing.

7. Experimental design is when, given the goal of a study, the simulator is used to choose an experiment which is most likely to allow these goals to be met. For example, if the outputs are temperatures at a number of locations on Earth, which of the locations should be observed to resolve enough uncertainty in the Earth's temperature to make the required decision.

The inferential tasks above usually require a large number of simulator evaluations. Analysis using complex computer models is therefore complicated when the simulators are expensive to evaluate. Some simulators can take months to run and the input space may be high dimensional. Therefore, an expensive simulator limits an analyst's ability to perform the tasks required as the output is unknown, apart from at the locations at which it has been evaluated.

4.1.1 Sources of Uncertainty and General Form

In order to make inferences about physical phenomena using complex computer models, the sources of uncertainty must be i) defined, ii) assigned random quantities and iii) organised into a model. Although the systems in question may vary greatly, many of the

uncertainties take analogous forms. Therefore, it is useful to have a general form when analysing complex computer models.

y is a multivariate random field indexed with respect to topological space $\mathcal{T}$. System values that cause system behaviour are described by multivariate random field S. The function that relates system values to system behaviour is $\mathcal{F}$. Therefore, the simplest relationship between system behaviour and system values is

$$y(t) = \mathcal{F}_{(t)}(S).$$

If S is continuous, it must undergo a discretisation process in order to be formulated as an input to $\mathcal{F}$. The discretisation is described by

$$x = \mathcal{D}(S)$$

where $\mathcal{D}$ is a function. The model becomes

$$y(t) = \mathcal{F}_{(t)}(x).$$

The relationship between the simulator and system behaviour may not be an equality but a function of a number of carefully chosen input-output combinations. For example, if the system behaviour of interest is an X-ray image and the simulator models a pencil beam directed at the detector through entity S, the system behaviour may be modelled as a summation of pencil beam simulations of simpler entities. This is particularly important when the input to the simulator is insufficient to describe S carefully enough to accurately mimic y. Instead of modelling the system behaviour as equal to the simulator evaluated at the system values, it is modelled as a function of a set of carefully chosen outputs of a set of simulations evaluated at carefully chosen inputs. The set of simulator evaluations required to describe system behaviour, notated $\mathcal{F}^r$, is given by

$$\mathcal{F}^r(X^r) = \left\{\mathcal{F}_{(t^{(i)})}(x^{r(i)})\right\}_{i=1:\tilde{\mathbb{N}}}$$

where $\mathcal{D}'(x) = X^r = \{x^{r(i)}\}_{i=1:\tilde{\mathbb{N}}}$ is the set of inputs at which the simulator must be evaluated to describe system behaviour. Then the system behaviour and simulator are related via

$$y(t) = \mathcal{H}_{(t)}(\mathcal{F}^r(X^r))$$

where $\mathcal{H}$ is a function of $\mathcal{F}^r(X^r)$. In summary, x is a discretisation of S which is well enough resolved to describe y. The input to the simulator x^r is a lower resolution version

of x. The system behaviour is then modelled as a function of a set of carefully chosen input-output combinations $\mathcal{F}^r(X^r)$. How an emulator of this form may be constructed will be discussed in more detail in Section 4.2.3 and exemplified using an industrial application in Section 4.3.

A simulator is a mathematical construct which is built to be informative for reality. However it is a model, and like all models it will be an imperfect representation of the physical system. This imperfection is referred to as model discrepancy and it is included in inferences using

$$y(t) = \mathcal{H}_{(t)}(\mathcal{F}^r(x^r)) + \epsilon_{md}(t) \tag{4.1.1}$$

where ϵ_{md} is a random vector describing beliefs about the difference between a simulator evaluated at the system values and the system behaviour. Structure is added to equation 4.1.1 to encode beliefs about how the model might differ from reality, see Chapter 6. For example, ϵ_{md} is assumed to be expectation zero and independent of all other random quantities. There may be observations of system behaviour available which will usually be measured imprecisely. In this thesis, the observation is modelled as

$$z(t) = y(t) \oplus \epsilon_{me}(t) \tag{4.1.2}$$

where ϵ_{me} is an expectation zero random field uncorrelated with $y(t)$. However, other noise models may be used instead, for example a multiplicative noise model. This construction relates a simulator, observations, system behaviour and system values in a general setting. Therefore, there is uncertainty about a system because there is uncertainty about

$$\{S, \mathcal{F}, \mathcal{H}, \mathcal{D}, \mathcal{D}', \epsilon_{md}, \epsilon_{me}\}. \tag{4.1.3}$$

Specifying these uncertainties may be challenging but this framework enables a process linking physical quantities and the abstract simulator. In this chapter, the focus will be on $\mathcal{F}$ and $\mathcal{H}$. The uncertainty in $\mathcal{F}$ will be described using an emulator and $\mathcal{H}$ will be specified directly using a hierarchical-superposition emulation strategy. In the next chapter, making inferences about x will be considered and in Chapter 6, the focus will be on ϵ_{md}.

4.2 Emulators

First, uncertainty in $\mathcal{F}$ is considered. If a simulator has a large input dimension and/or it is expensive to evaluate, it is appropriate to consider the output of the simulator as

uncertain everywhere, apart from the locations at which it has been evaluated. Beliefs about the simulator are organised as a multivariate random field and evaluations of a deterministic simulator are locations on that field. Then inferences about the simulator at any location may be made using the methodology for adjusting beliefs about multivariate random fields by vectors the elements of which are locations on a multivariate random field, see Chapter 2.

An emulator is defined as a statistical model of a simulator. The term surrogate or meta-model is also used to describe models of simulators although they are more often used to describe a fast approximation without the corresponding uncertainty judgements. Emulators are usually faster to evaluate so enable inference in a reasonable time frame. Emulators have been used to quantify uncertainty about complex computer models in many applications [23, 24, 32, 34–39, 41, 42, 52]. The approaches mostly use a full Bayes approach employing Gaussian processes (as in [31]) or the Bayes linear approach (as in [22]). The Managing Uncertainty in Complex Models (MUCM) website [44] is an excellent resource which contains a toolkit and a large amount of relevant literature.

In this work, an emulator is built using the form

$$\mathcal{F}_{(t)}(x) = \sum_{m=1}^{r} \beta_{(m,t)} h_{(m)}(x) \oplus \mathcal{W}_{(t)}(x).$$

A set of evaluations of the simulator are used to adjust beliefs about the simulator output everywhere. This is known as the emulator training-set, notated $D^{\mathcal{F}} = \{\mathcal{F}(x^{(i)})\}_{i=1:\mathbb{N}}$. The set of inputs used to create $D^{\mathcal{F}}$ is notated $X = \{x^{(i)}\}_{i=1:\mathbb{N}}$. Adjusted beliefs are then calculated using Section 2.2.2 or approximated using GLS estimates, key-formula 2.2.1.

Toy Example: Deterministic Emulator

To illustrate emulating a deterministic simulator, a toy example is introduced. The toy simulator is given by

$$\mathcal{F}(x) = 1 + 0.5x + 0.1\sin(10x).$$

Figure 4.1 shows an emulator for the toy simulator. The mean function was specified using $h(x) = (1, x)$ and the residual process correlation structure was specified using equation 2.2.8 with correlation length $l = 0.3$ and $p = 2$. The GLS approach, see key-formula 2.2.1, is used to calculate the approximate adjusted beliefs. The simulator output is always within the three sigma credible interval and the emulator interpolates the training points where the simulator output is known.

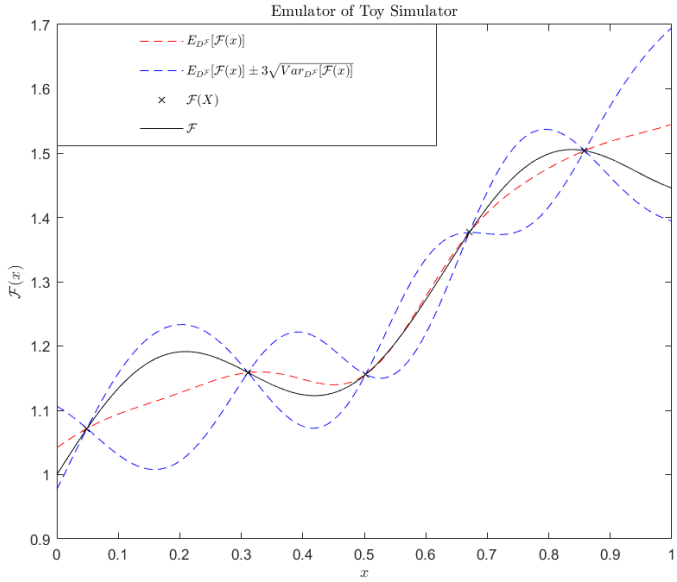

Figure 4.1: Adjusted beliefs about the toy simulator.

4.2.1 Emulating Stochastic Simulators

A stochastic simulator is when each evaluation at the same input is not necessarily the same. It is assumed that evaluations at the same input are second order exchangeable and, by the representation theorem (see equation 3.2.1), may be written in the form

$$\mathcal{F}^{(i)}_{(t)}(x) = M(\mathcal{F}_{(t)}(x)) \oplus R^{(i)}(\mathcal{F}_{(t)}(x))$$

where $M(\mathcal{F}_{(t)}(x))$ describes simulator's population mean and $R^{(i)}(\mathcal{F}_{(t)}(x))$ is the residual from this mean. Given that it is asserted that the set $\{V^{(1)}_{\mathcal{F}(t)}(x), \cdots\}$ is second order exchangeable, $V^{(i)}_{\mathcal{F}(t)}(x)$ may be written as

$$V^{(i)}_{\mathcal{F}(t)}(x) = R^{(i)}(\mathcal{F}_{(t)}(x))R^{(i)}(\mathcal{F}_{(t)}(x))^T = M(V_{\mathcal{F}(t)}(x)) \oplus R^{(i)}(V_{\mathcal{F}(t)}(x))$$

where $M(V_{\mathcal{F}(t)}(x))$ is the simulator's population variance. Then a multivariate random field is used to model both the simulator's population mean and its variance, given by

$$M(\mathcal{F}_{(t)}(x)) = \mathcal{M}_{M(t)}(x) \oplus \mathcal{W}_{M(t)}(x) \text{ and}$$
$$M(V_{\mathcal{F}(t)}(x)) = \mathcal{M}_{V(t)}(x) \oplus \mathcal{W}_{V(t)}(x).$$

In this work, two methods of emulating stochastic simulators were tested. The first is a full Bayes linear two stage update which i) makes repeat simulations at a number of locations, ii) adjust beliefs about the simulator's population variance and iii) uses the adjusted expectation of the simulator's population variance as a point estimate in the adjustment

of its population mean. See Appendix A.6 for details. This approach was effective and used in early versions of the algorithm. However, as the stochastic nature of the simulator is Poisson noise, the population variance can be assumed to be proportional to the population mean. Furthermore, the simulator's population variance was only of interest to the extent that knowledge of its value would reduce uncertainty in the simulator's population mean. For the applications presented, testing showed that the direct specification method presented below was adequate to achieve the goals of this study and computationally more efficient.

Direct Specification

If repeated simulations at the same input are not available, a direct specification approach can be adopted. This may be combined with a separable covariance formulation (see equation 2.2.7) to create an efficient approximation of adjusted beliefs. In this approach,

$$w(x,t) = \frac{M(V_{\mathcal{F}(t)}(x))}{\mathbb{V}ar[\mathcal{W}_{M(t)}(x)]} \tag{4.2.1}$$

is specified directly. If it can be assumed that $w(x,t)$ is invariant to t, the GLS estimates from key formula 2.2.1 may be employed. An alternative tractable form is when it can be assumed that the population variance of the simulator is independent of the inputs given the population mean of the simulator. Then $w(x,t)$ may be specified as a function with scalar input

$$w(x,t) = \mathcal{Q}(M(\mathcal{F}_{(t)}(x))) \approx \mathcal{Q}(\mathcal{F}_{(t)}(x))$$

where $\mathcal{Q}$ is a function specified directly and tested using validation procedures. For example, a stochastic simulator of an X-ray examination procedure can be viewed as an integral over Poisson processes. Therefore, it is reasonable to assume that the simulator's population variance is proportional to simulator's population mean. This is a simple procedure with a number of approximations. However, if the interest is in the simulator's population mean, it can validate well and has computational advantages over the full Bayes linear update.

In general, we advocate starting with the two stage Bayes linear approach in Appendix A.6. Then, if the computational requirements inhibit inference in a useful time-frame, the direct specification methods can be used if validation shows the approximations made are not having too large an impact on the inference.

4.2.2 Look-up-tables

Although emulators are usually relatively cheap to evaluate, it is not a trivial calculation. The cost of an evaluation of an emulator scales with $O(\mathbb{N}^3)$ and, for "large" $\mathbb{N}$ (over 10000), if an inferential procedure requires a "very large" (over 10000000) number of emulator evaluations, the process becomes computationally challenging again. A common approach to dealing with this issue is to perform the analysis in waves. The emulator at each wave only has a "small" (under 10000) number of training points. Each wave focuses the region the emulator represents and creates a new training-set in the refocused area of interest. See, for example, [32, 35–37]. This resolves the issue but it requires additional simulations. If the aim is to create an algorithm to perform the inference in real-time and the simulator is expensive to evaluate, running the simulator online is not feasible.

An alternative is to produce all the simulations required to resolve sufficient uncertainty in the simulator offline and perform a wave based approach. Many emulators have to be built, validated and evaluated in an online algorithm. This is problematic if the aim is to provide robust software to non-statistical experts. It would be better to be able to have all the information from the simulator evaluations included in a single process, which can be validated offline and evaluated in a reasonable time-frame.

Here instead of using the emulator, a look-up-table of the emulator is used. A look-up-table is a process which replaces the emulator with a simpler operation. They are often used in computer science to replace run-time calculations with offline processing time. The look-up-table consists of a large number of emulator evaluations which are chosen to fill the space of interest. The look-up process is then a weighted average of the 2^p closest points, where the weights are calculated by linear interpolation and p is the dimension of the input. Look-ups are an embarrassingly paralisable problem meaning that the computation time can be decreased using multi-threading techniques where the operations are shared by many processors. This implies that Graphics Processing Units (GPUs) may be used to significantly decrease the computation time of the look-up process. The cost is that the look-up-table has to be calculated offline and an error between an evaluation of the look-up-table and the emulator is incurred.

In order to use look-up-tables for emulation, both the expectation and variance of the emulator require a look-up-table. Then adjusted beliefs about a simulator with scalar

output are approximated using

$$\mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}(x)] + \epsilon_{\mathcal{LU}} \approx \mathcal{LU}^{\mathbb{E}}(x) \text{ and}$$
$$\mathbb{V}ar[\mathcal{F}(x) - \mathcal{LU}^{\mathbb{E}}(x)] \approx \mathcal{LU}^{\mathbb{V}ar}(x) + \mathbb{V}ar[\epsilon_{\mathcal{LU}}]$$

where $\mathcal{LU}^{\mathbb{E}}$ is the function describing the look-up process of the adjusted expectations, $\mathcal{LU}^{\mathbb{V}ar}$ is the function describing the look-up process of the adjusted variances and $\epsilon_{\mathcal{LU}}$ is the additional error induced by the look-up process. $\mathbb{V}ar[\epsilon_{\mathcal{LU}}]$ may be approximated as the sample variance of a set of residuals between emulator evaluations and look-up values.

Look-up tables can make inference in a useful time-frame possible, however, there are limitations. The variance of $\epsilon_{\mathcal{LU}}$ and the cost of each look-up increases rapidly with the dimension of x. Therefore, calculating the adjusted covariance using a look-up table is difficult as the dimension of the look-up table scales with p^2. The net effect is that more memory has to be assigned to the adjusted covariance table making the adjusted expectation and variance tables less accurate. This is not a problem if only adjusted expectation and variance is used in the subsequent analysis. However, if an emulator is built using equation 4.1.1, the uncertainty estimation can be hampered by not being able to calculate adjusted covaraiances. How this affects the application presented in this thesis is discussed in more detail in Section 4.3.5.

The use of look-up-tables poses interesting questions about how best to design training-sets of simulations and how different types of multiple output emulators may be manipulated to assimilate all available evaluations of the simulator. Suppose that the aim is to build a look-up-table for a scalar simulator and the number of evaluations of $\mathcal{F}$ available means that the covariance matrix is too large to hold in memory. The input x is split into two vectors (x_a, x_b). The set of points in the look-up-table is notated $X^{\mathcal{LU}}$ and may be written in the form

$$\mathrm{X}^{\mathcal{LU}} = \begin{bmatrix} (x_a^{(1)}, x_b^{(1)}) & \dots & (x_a^{(1)}, x_b^{(d)}) \\ \vdots & \ddots & \vdots \\ (x_a^{(c)}, x_b^{(1)}) & \dots & (x_a^{(c)}, x_b^{(d)}) \end{bmatrix}.$$

The training simulations are designed to have the same structure as $X^{\mathcal{LU}}$ in that for every $x_a^{(i)}$, the set of x_bs is consistent. Then the emulator is built in two stages.

1. For every $x_a^{(i)}$ in $D^{\mathcal{F}}$, an emulator of

$$\mathcal{F}(x_a^{(i)}, x_b)$$

is built and the adjusted beliefs are calculated for every value in $\{x_b^{(i)}\}_{i=1:d}$.

2. An emulator is then built over x_a as a multiple output stochastic emulator of the form

$$\mathcal{F}'(x_a^{(i)}) = \left(\mathbb{E}_{D^{\mathcal{F}(i)}}[\mathcal{F}(x_a^{(i)}, x_b^{(j)})]\right)_{j=1:d}$$

where $D^{\mathcal{F}(i)}$ is the subset of $D^{\mathcal{F}}$ where $x_a = x_a^{(i)}$. Finally, the emulator of $\mathcal{F}'$ is evaluated for every value in $\{x_a^{(i)}\}_{i=1:c}$ to create the look-up-tables.

This approach is efficient when it is reasonable to assume $\mathcal{F}'$ has a separable covariance structure, see equation 2.2.7, and the stochastic component may be specified directly, as in equation 4.2.1, where $w(x,t)$ is invariant to t.

4.2.3 Hierarchical-superposition Emulators

The novel strategy for emulating high dimensional simulators of imaging systems with random field inputs and outputs will now be presented. Doing the analysis with $\mathcal{F}$ directly may not be possible because: i) x may be too high dimensional and too expensive to emulate directly or ii) x may be too simple to predict system behaviour for arbitrary S. In the second case, the analysis should be viewed as adjusting beliefs about more complex simulator $\mathcal{F}'$ by $\mathcal{F}$. This idea is not unusual in Bayesian emulation: [27] and [34] use evaluations of a lower resolution simulators to make inferences for a higher resolution one and [23] uses the simulator available to make inferences for a hypothetical improved one.

To motivate the approach, consider an X-ray source producing a cone of photons which intersects an entity and then a detector. Each photon produced has a direction and an energy. The output recorded at the detector is the sum of the energy deposited by photons hitting the detector at a given location. This can be viewed as an integral of energy deposited at the detector over initial directions and initial energies. The aim of the simulator is to calculate this integral. One way of achieving this goal is to decompose the output into the effect of individual photons. This is how the simulator in the following section calculates the output. A photon is sampled with an initial direction and energy. It is then tracked until it interacts with the detector or is absorbed. Although an accurate method, it is slow. An alternative way of approximating this integral is to split up the initial directions into a set which hit the centre of every pixel on the detector and sum over that discrete set. This method can be made more efficient for two reasons: a) the dependence on the entity can be assumed to be dependent on a small part of the entity and b) the initial direction is irrelevant given the relative position of i) source, ii) entity

and iii) detector is fixed. An emulator is built of the response of the simulator to a single set of directions and the integral is performed by summing over adjusted expectations.

This approach was inspired by the use of Point Spread Functions (PSFs) to model imaging systems. A PSF describes the response of an imaging system to a point source and is often described as the impulse response of an optical system. The process of producing an image is then modelled as the superposition of the imaging system's response to a point source stimulus

$$\mathcal{F}_{(t)}(S) = \int_{t'} O(t')\mathcal{IRF}(|t - t'|)dt' \tag{4.2.2}$$

where $O(t') = \mathcal{F}^*(S(t))$, referred to as the object plane, is the quantity about which inferences will be made and $\mathcal{IRF}$ is the impulse response function. The advantage of this approach is that the complex imaging system is decomposed into two simple components: i) a lower dimensional function $\mathcal{F}^*$ which models the process of producing the object plane from an entity and ii) a function with a scalar input $\mathcal{IRF}(|t-t'|)$. So instead of modelling the process of making an image as a complex function of high dimensional S, the process is modelled as the convolution of a low dimensional simulator and an impulse response function. Given that models with analogous forms to equation 4.2.2 have been useful in modelling physical phenomena, it is reasonable to assume that they would also be useful in modelling a simulator of that phenomena.

The remainder of this section will propose how to use signal processing modelling techniques within the Bayesian emulation framework and present methods for relaxing the assumptions implicit in equation 4.2.2. Instead of emulating a simulator with high dimensional inputs and outputs directly, models of lower dimensional components will be built and combined with products, summations and catenations. Given that the function of lower dimensional components will only approximate the simulator in question, residual processes are added to resolve additional uncertainty and quantify the remaining unresolved portion.

The method proposed assumes the full simulator may be decomposed into a number of simulations which model the response of the system to a stimulus acting at a location on a random field. For example, the stimulus could be a pencil beam of X-rays directed at a location on a detector through an entity or a pump being activated on an oil field. It is assumed that the simulator of interest is described by

$$\mathcal{F}(x, M = 0)$$

where M defines the mode of the simulator. A mode is defined as an input that can result in a non-continuous change in the simulator output. x and $\mathcal{F}(x, M = 0)$ are assumed to be multivariate fields with respect to the same topological space $\mathcal{T}$, meaning that they may be written in the form

$$\begin{aligned}\mathcal{F}(x, M) &= (\mathcal{F}_{(t^{(1)})}(x, M), \cdots, \mathcal{F}_{(t^{(n)})}(x, M)) \text{ and} \\ x &= (x(t^{(1)}), \cdots, x(t^{(n)})) \text{ where} \\ x(t) &= (x_{(1)}(t), \cdots, x_{(p)}(t))\end{aligned}$$

and $t^{(i)} \in \mathcal{T}$. It is assumed that the response of the simulator to a stimulus acting at location $t^{(i)}$ may be viewed by changing the mode of the simulator and the simulator run at mode $M = 0$ may be written as the superposition of the simulator evaluated at different modes. This assumption is written as

$$\mathcal{F}_{(t)}(x, M = 0) = \sum_{i=1}^{n} \mathcal{F}_{(t)}(x, M = t^{(i)}) + \epsilon \tag{4.2.3}$$

where $M = t^{(i)}$ is a stimulus acting at location $t^{(i)}$ and ϵ is small enough to be ignored. Then instead of building a statistical model for $\mathcal{F}_{(t)}(x, M = 0)$ directly, $\mathcal{F}_{(t)}(x, M \neq 0)$ is emulated. It is preferable to emulate $\mathcal{F}_{(t)}(x, M \neq 0)$ when i) assumptions may be made to reduce its dimension ii) it is smoother iii) it is less expensive to evaluate or iv) more informed prior knowledge about it is available. A number of assumptions that may be made about simulators of this form will now be presented with accompanying methodology for how they may be emulated.

Dirac Assumption

The simplest assumption is when the output of the simulator at any location is only dependent on a small subset of the inputs.

$$\begin{aligned}\mathcal{F}_{(t)}(x, M = t^{(i)}) &= 0 \text{ for } t \neq t^{(i)} \text{ and} \\ \mathcal{F}_{(t)}(x, M = t) &= \mathcal{F}_{(t)}(\mathcal{G}(x_a(t)), M = t) \oplus \mathcal{W}^a(x_a(t)).\end{aligned} \tag{4.2.4}$$

Here $\mathcal{G}(x_a(t))$ is a function which returns a vector the same size as x with a subset equal to $x_a(t)$. $x_a(t)$ is referred to as the active inputs for mode t. The residual process has two roles: capture variation correlated with $x_a(t)$ and the variation induced by the remaining inactive inputs. Then using equations 4.2.4 and 4.2.3, the simulator output may be written

$$\mathcal{F}(x, M = 0) = \left(\mathcal{F}_{(t)}(\mathcal{G}(x_a(t^{(i)})), M = t) \oplus \mathcal{W}^a(x_a(t^{(i)}))\right)_{i=1:n}. \tag{4.2.5}$$

Here $\mathcal{F}_{(t)}(\mathcal{G}(x_a(t)), M = t)$ is referred to as the primitive simulator and the adjusted expectation of the full simulator output may be inferred from an emulator of the form

$$\mathcal{F}_{(t)}(\mathcal{G}(x_a(t)), M = t) = \mathcal{M}(x_a(t)) \oplus \mathcal{W}(x_a(t)). \tag{4.2.6}$$

Then an emulator with input dimension pn and output dimension n may be built using a scalar emulator with input $x_a(t)$. The emulator training-set is defined by $D^{\mathcal{F}} = \{\mathcal{F}_{(t)}(\mathcal{G}(x_a^{(i)}(t)), M = t)\}_{i=1:\mathbb{N}}$ and the adjusted beliefs are given by

$$\mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}_{(t)}(x, M = 0)] = \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}_{(t)}(\mathcal{G}(x_a(t)), M = t)] \text{ and} \tag{4.2.7}$$

$$\mathbb{V}ar_{D^{\mathcal{F}}}[\mathcal{F}_{(t)}(x, M = 0)] = \mathbb{V}ar_{D^{\mathcal{F}}}[\mathcal{F}_{(t)}(\mathcal{G}(x_a(t)), M = t)] + \mathbb{V}ar[\mathcal{W}^a(x_a(t))]. \tag{4.2.8}$$

Beliefs about the residual process $\mathcal{W}^a$ may be adjusted using a number of simulations with $M = 0$. The set of simulations used to adjust beliefs about the residual process is given by $\left\{\mathcal{F}(x^{(i)}, M = 0)\right\}_{i=1:\tilde{\mathbb{N}}}$. For each of the $\tilde{\mathbb{N}}$ simulations, the residual between the simulator output and the adjusted expectation of the simulator output is calculated using

$$e^{(i)} = \mathcal{F}(x^{(i)}, M = 0) - \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}(x^{(i)}, M = 0)]$$

where $\mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}(x^{(i)}, M = 0)]$ is calculated by catenating n adjusted expectations using equations 4.2.5 and 4.2.7. Then the set

$$\begin{aligned}
&\{e^{(i)}(t^{(j)})\}_{i=1:\tilde{\mathbb{N}},j=1:n} \text{ where} \\
e^{(i)}(t^{(j)}) &= \mathcal{F}_{(t^{(j)})}(x^{(i)}, M = 0) - \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}_{(t)}(\mathcal{G}(x_a^{(i)}(t^{(j)})), M = t)] \\
&= \mathcal{W}^a(x_a^{(i)}(t^{(j)})) \oplus \mathbb{A}_{D^{\mathcal{F}}}[\mathcal{F}_{(t)}(\mathcal{G}(x_a^{(i)}(t^{(j)})), M = t)]
\end{aligned}$$

may be used to adjust beliefs about the residual process. Using the properties of adjusted expectations, variances and covariances set out in Section 2.1.1, the second order specification is given by

$$\begin{aligned}
\mathbb{E}[e^{(i)}(t^{(j)})] &= 0, \\
\mathbb{V}ar[e^{(i)}(t^{(j)})] &= \mathbb{V}ar[\mathcal{W}^a(x_a^{(i)}(t^{(j)}))] + \mathbb{V}ar_{D^{\mathcal{F}}}[\mathcal{F}_{(t)}(\mathcal{G}(x_a^{(i)}(t^{(j)})), M = t)], \\
\mathbb{C}ov[e^{(i)}(t^{(j)}), e^{(i')}(t^{(j')})] &= \mathbb{C}ov[\mathcal{W}^a(x_a^{(i)}(t^{(j)})), \mathcal{W}^a(x_a^{(i')}(t^{(j')}))] + \\
&\quad \mathbb{C}ov_{D^{\mathcal{F}}}[\mathcal{F}_{(t)}(\mathcal{G}(x_a^{(i)}(t^{(j)})), M = t), \mathcal{F}_{(t)}(\mathcal{G}(x_a^{(i')}(t^{(j')})), M = t)], \\
\mathbb{C}ov[e^{(i)}(t^{(j)}), \mathcal{W}^a(x_a^{(i')}(t^{(j')}))] &= \mathbb{C}ov[\mathcal{W}^a(x_a^{(i)}(t^{(j)})), \mathcal{W}^a(x_a^{(i')}(t^{(j')}))].
\end{aligned}$$

The aim is to build an emulator of a simulator with, for example, 1000 inputs. Building this emulator directly would be difficult using conventional emulation procedures so a simplifying assumption is required. Given that the assumption is known not to be true,

a residual process is added to encode beliefs about the error induced by the assumption. If $\mathcal{F}(x, M = 0)$ is available, beliefs about the residual process may be adjusted using evaluations of that simulator. This process uses ideas from signal processing to resolve the majority of the variation and Bayesian emulation by encoding beliefs about the error induced by the simplifying assumption in a residual process. If the residual process variance is too large for inferences to be useful, the assumptions made need to be relaxed.

Stationary-Isotropic Assumption

The next assumption considered is when it is assumed that

$$\mathcal{F}_{(t)}(x, M = t^{(i)}) = \mathcal{F}_{(t)}(\mathcal{G}(x_a(t^{(i)})), M = t)\mathcal{IRF}(|t - t^{(i)}|) \oplus \mathcal{W}^b(x_a(t^{(i)}), |t - t^{(i)}|)$$

where $\mathcal{IRF}(|t-t^{(i)}|)$ is referred to as an impulse response function. The impulse response function may be inferred from any simulation $\mathcal{F}(\mathcal{G}(x_a(t)), M = t)$. $\mathcal{F}_{(t)}(\mathcal{G}(x_a(t)), M = t)$ is still the primitive simulator which may be emulated using equation 4.2.6. The impulse response function is a convenient simplification as efficiency can be significantly improved using the fast Fourier transform.

The residual process $\mathcal{W}^b$ accounts for the remaining variation. To adjust beliefs about the residual process a set of simulations, notated $\{\mathcal{F}(x^{(i)}, M = t)\}_{i=1:\tilde{\mathrm{N}}}$, may be used. Then the set

$$\{e^{(i)}(t^{(j)})\}_{i=1:\tilde{\mathrm{N}},j=1:n} \text{ where}$$
$$e^{(i)}(t^{(j)}) = \mathcal{F}_{(t^{(j)})}(x^{(i)}, M = t) - \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}_{(t)}(\mathcal{G}(x_a^{(i)}(t)), M = t)]\mathcal{IRF}(|t - t^{(j)}|)$$

may be used to adjust beliefs about $\mathcal{W}^b$.

Non-Stationary-Isotropic Assumption

The next logical extension is to make the impulse response function dependent on $x_a(t)$. This assumption is given by

$$\begin{aligned}\mathcal{F}_{(t)}(x, M = t^{(i)}) = \mathcal{F}_{(t)}(\mathcal{G}(x_a(t^{(i)})), M = t)\mathcal{IRF}(x_a(t^{(i)}), |t - t^{(i)}|) \\ \oplus \mathcal{W}^c(x_a(t^{(i)}), |t - t^{(i)}|).\end{aligned} \tag{4.2.9}$$

where the impulse response function may be calculated using $\mathcal{IRF}(x_a(t^{(i)}), |t - t^{(i)}|) = \frac{\mathcal{F}_{(t^{(i)})}(\mathcal{G}(x_a(t^{(i)})), M=t)}{\mathcal{F}_{(t)}(\mathcal{G}(x_a(t^{(i)})), M=t)}$. Therefore, equation 4.2.9 may be rewritten in the form

$$\mathcal{F}_{(t)}(x, M = t^{(i)}) = \mathcal{F}_{(t)}(\mathcal{G}(x_a(t^{(i)})), M = t^{(i)}) \oplus \mathcal{W}^c(x_a(t^{(i)}), |t - t^{(i)}|).$$

The disadvantage of this approach is that the computational advantages of the fast Fourier transform are lost but, if making this assumption resolves sufficient uncertainty in the simulator to make useful inferences, it can be made efficient using look-up-tables. In this case, the primitive simulator is $\mathcal{F}(\mathcal{G}(x_a(t)), M = t)$, which may be emulated using

$$\mathcal{F}_{(t^{(i)})}(\mathcal{G}(x_a(t)), M = t) = \mathcal{M}(x_a(t), |t - t^{(i)}|) \oplus \mathcal{W}(x_a(t), |t - t^{(i)}|).$$

The emulator training-set is now $D^{\mathcal{F}} = \{\mathcal{F}(\mathcal{G}(x_a^{(i)}(t)), M = t)\}_{i=1:\mathbb{N}}$ and beliefs about the residual process may be adjusted using a set of simulations $\{\mathcal{F}(x^{(i)}, M = t)\}_{i=1:\tilde{\mathbb{N}}}$. Then the set

$$\begin{aligned}&\{e^{(i)}(t^{(j)})\}_{i=1:\tilde{\mathbb{N}},j=1:n} \text{ where}\\ &e^{(i)}(t^{(j)}) = \mathcal{F}_{(t^{(j)})}(x^{(i)}, M = t) - \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}_{(t^{(j)})}(\mathcal{G}(x_a^{(i)}(t)), M = t)])\end{aligned}$$

may be used to adjust beliefs about $\mathcal{W}^c$. Using the properties of adjusted expectations, variances and covariances set out in Section 2.1.1, the second order specification is given by

$$\begin{aligned}\mathbb{E}[e^{(i)}(t^{(j)})] &= 0,\\ \mathbb{V}ar[e^{(i)}(t^{(j)})] &= \mathbb{V}ar[\mathcal{W}^c(x_a^{(i)}(t), |t - t^{(j)}|)]\\ &\quad + \mathbb{V}ar_{D^{\mathcal{F}}}[\mathcal{F}_{(t^{(j)})}(\mathcal{G}(x_a^{(i)}(t)), M = t)],\\ \mathbb{C}ov[e^{(i)}(t^{(j)}), e^{(l)}(t^{(k)})] &= \mathbb{C}ov[\mathcal{W}^c(x_a^{(i)}(t), |t - t^{(j)}|), \mathcal{W}^c(x_a^{(l)}(t), |t - t^{(k)}|)] +\\ &\quad \mathbb{C}ov_{D^{\mathcal{F}}}[\mathcal{F}_{(t^{(j)})}(\mathcal{G}(x_a^{(i)}(t)), M = t), \mathcal{F}_{(t^{(k)})}(\mathcal{G}(x_a^{(l)}(t)), M = t)],\\ \mathbb{C}ov[e^{(i)}(t^{(j)}), \mathcal{W}^c(x_a(t^{(l)}), |t^{(l)} - t^{(k)}|)] &= \mathbb{C}ov[\mathcal{W}^c(x_a^{(i)}(t), |t - t^{(j)}|), \mathcal{W}^c(x_a(t^{(l)}), |t^{(l)} - t^{(k)}|)].\end{aligned}$$

Non-Stationary-Non-Isotropic Assumption

The next extension is to consider when there is still significant dependence on x which can be parameterised into a concise form. The assumption is given by

$$\mathcal{F}_{(t^{(i)})}(x, M = t^{(j)}) = \mathcal{F}_{(t^{(i)})}(\mathcal{G}(x_a(t^{(j)}), u^{(i,j)}), M = t^{(j)}) \oplus \mathcal{W}^d(x_a(t^{(j)}), u^{(i,j)}, |t^{(i)} - t^{(j)}|)$$

where the vector $u^{(i,j)} = \mathcal{RT}(x, t^{(i)}, t^{(j)})$. $\mathcal{RT}(x, t^{(i)}, t^{(j)})$ is a known function of the inputs, for example it could be an average of all locations in x_a on a line between $t^{(i)}$ and $t^{(j)}$. $u^{(i,j)}$ is not a subset of the input vector so is not considered an active input. It parameterises the remaining dependence on the x in a concise way. For example, if $\mathcal{F}$ models the amount of carbon dioxide in the atmosphere and $x_a(t^{(i)})$ is the amount produced at location $t^{(i)}$, $u^{(i,j)}$

could be the average carbon dioxide absorption between $t^{(i)}$ and $t^{(j)}$. This assumption is key to the approach used in this thesis. It is essentially saying that a function can be designed such that if

$$\mathcal{TR}(x', t^{(i)}, t^{(j)}) = \mathcal{TR}(x, t^{(i)}, t^{(j)}),$$

then it is reasonable to assert that

$$\mathcal{F}_{(t^{(i)})}(x', M = t^{(j)}) = \mathcal{F}_{(t^{(i)})}(x, M = t^{(j)}) \oplus \mathcal{W}^d(x_a(t^{(j)}), u^{(i,j)}, |t^{(i)} - t^{(j)}|).$$

Then for any x, there exists a function $\mathcal{G}$

$$\mathcal{TR}(x, t^{(i)}, t^{(j)}) = \mathcal{TR}(\mathcal{G}(x_a(t^{(i)}), u^{(i,j)}), t^{(i)}, t^{(j)}).$$

The primitive simulator in this case is emulated using

$$\mathcal{F}_{(t^{(i)})}(\mathcal{G}(x_a(t^{(j)}), u^{(i,j)}), M = t^{(j)}) = \mathcal{M}(x_a(t^{(j)}), u^{(i,j)}, |t^{(i)} - t^{(j)}|) \oplus \mathcal{W}(x_a(t^{(j)}), u^{(i,j)}, |t^{(i)} - t^{(j)}|).$$

The emulator training-set is $D^{\mathcal{F}} = \{\mathcal{F}(\mathcal{G}(x_a^{(i)}(t), u^{(i)}), M = t)\}_{i=1:\mathbb{N}}$. The set of reduced inputs to the simulator is given by $\mathcal{D}'(x) = \{x_a(t^{(i)}), u^{(i,j)}\}_{i=1:n,j=1:n}$ and the residual process may be adjusted using the set

$\{e^{(j)}(t^{(i)})\}_{i=1:\tilde{\mathbb{N}},j=1:n}$ where
$e^{(j)}(t^{(i)}) = \mathcal{F}_{(t^{(i)})}(x^{(j)}, M = t^{(k)}) - \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}_{(t^{(i)})}(\mathcal{G}(x_a^{(j)}(t^{(k)}), u^{(i,j)}), M = t^{(k)})]$.

The superposition principle is useful for emulation for the same reason that it is useful for modelling physical phenomena. It is an approximation which enables the majority of the variation to be efficiently resolved. Residual processes then have an easier job in resolving the remainder of the variation requiring less simulations to resolve sufficient uncertainty in the simulator. The process of building an emulator is as follows: make a superposition assumption on the simulator output, employ an assumption analogous to those used in signal processing to resolve the majority of the variation and adjust beliefs about the residual process using a set of evaluations of the target simulator. The difficulty is in designing $\mathcal{G}$ such that the assumptions are reasonable. This is likely to be situation specific. How $\mathcal{G}$ may be constructed will be clarified in the example in the following section.

Given an expensive simulator with a very high dimensional input and output, the toolkit emulation approach [44] can perform badly as i) any reasonable number of simulations will only sparsely represent the input space and ii) the memory requirements for storing the training-set and evaluating the emulator are prohibitively large. In these

cases, a simplifying assumption is required to be able to build an emulator or dimension reduction techniques [26] have to be employed. The advantage of signal processing assumptions is that the resolution of the input and output is not altered. The advantage of the emulation approach over direct application of signal processing techniques is that i) the residual error is quantified and partially resolved, ii) the assumptions being made are clearly set out and validated and iii) assumptions which induce too large an error may be relaxed. The advantages of the hierarchical-superposition approach enable improvements, over pure signal processing techniques, for modelling a simulator of an X-ray examination system enable a solution to problem statement 1.1.2.

4.3 Predicting Scatter

It is now time to move onto the second industrial application, predicting the intensity in an X-ray image caused by scattered photons.

4.3.1 Preliminaries: X-ray Scattering

An X-ray image is produced by recording the energy deposited at a detector from photons directed through an entity. The photons can either i) traverse the object unperturbed, ii) be absorbed or iii) scatter. The intensity recorded by a pixel at location (i, j) is described by

$$Y(i,j) = Y^{Sc}(i,j) + Y^{Dr}(i,j)$$

where $Y^{Sc}(i,j)$ is the energy deposited at the detector from photons which have undergone a scattering interaction and $Y^{Dr}(i,j)$ from those which have not interacted with the entity. Y^{Sc} is referred to as the scattering component and Y^{Dr} is the direct component. $Y^{Sc}(i,j)$ is considered a form of spatially correlated error which degrades image quality by adding a deformed projection of the entity onto the image. When a diagnostic image is produced, Y is classified into levels and each level is assigned a shade of grey in the visual output. The addition of scatter means there are not enough grey levels to accurately represent all the necessary information. Scatter increases the range of values needed to display the entire image meaning that fine detail is lost when Y is categorised in the output shown to clinicians. This is visualised as vignetting and the loss of fine detail.

Figure 4.2 shows an image of a phantom designed to assess detector resolution before and after scatter correction is applied. The left image is saturated in the middle and at

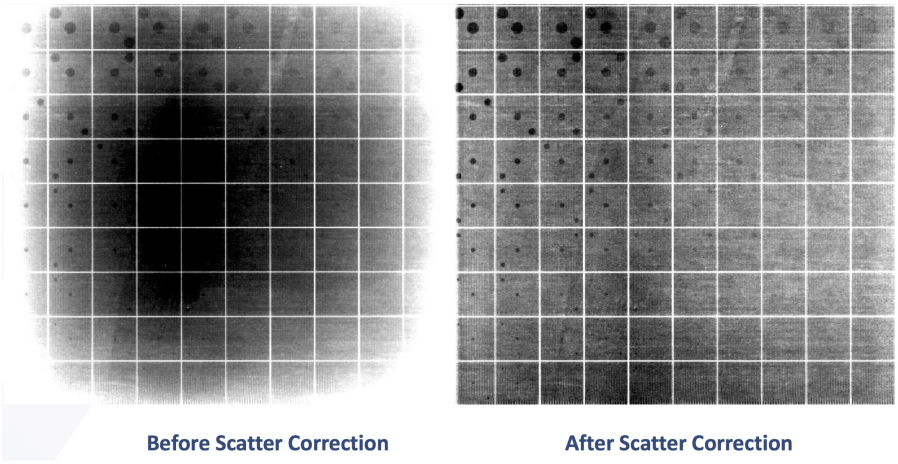

Figure 4.2: Left: an image without scatter correction. Right: an image with the scatter correction algorithm developed in this thesis applied.

the edges. In this context, saturated means that all grey levels have been used so only one colour, either white or black, is available for saturated regions. As shown by the small dots not being visible in the left hand image, fine detail is lost. The right image shows a scatter corrected image where a larger number of dots are visible and the vignetting is removed. If the small dots are what a clinician requires to make a diagnosis, the advantage of reducing the effect of scattered photons is clear.

It is also of interest to infer the patient's chemical composition. For example, areal Bone Mineral Density (aBMD) is useful to deduce bone health and diagnose osteoporosis. Scattered photons cause problems when making inferences about the patient's composition. The direct component at a pixel on the detector is only dependent on the entity composition in a small local area around the path between the source and that pixel. The scatter component is dependent on the chemical composition of the entire entity, which makes performing the inference required much more complex.

Currently the most common solution to the scattering problem is an anti-scatter-grid (ASG) . The ASG is an engineering solution which is designed to block photons which have undergone a scattering interaction. However, it also blocks direct beam photons, which increases the Poisson noise in the image. To restore the signal-to-noise-ratio required for diagnostic image quality, more photons are produced at the source. Increasing the number of photons raises the patient dose, which has an adverse effect on health. For effective scatter removal, the ASG must also be appropriately aligned (which is challenging for operators) otherwise oscillating artefacts appear in the image degrading image quality.

Finally, ASGs are not perfect and a significant proportion of scatter makes it through. This adversely effects image quality and the inferences made about an entity's composition. Figure 4.3 shows the same entity with and without an ASG. The grid has not been well enough aligned and there are vertical artefacts degrading image quality.

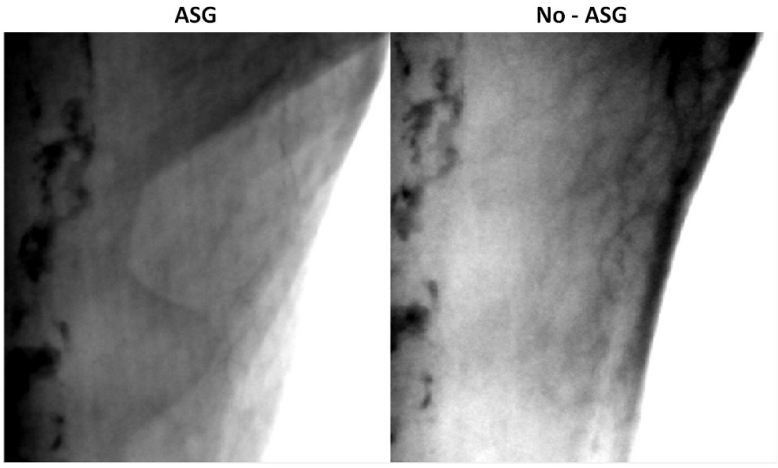

Figure 4.3: Left: an X-ray image of a cadaver with an ASG. Right: an X-ray image of the same cadaver with no ASG.

Given the problems with ASGs and the impact of scatter on image quality, there have already been a number of attempts at software based scatter removal. There are a wide range of proposed methods which in general try to calculate the expected scatter and remove it. [63] presents an image flattening technique, they de-trend a polynomial fit which is reported to be scatter. [65] presents a method of adding a high frequency oscillating structure, similar to a Multi Adsorption Plate (MAP), to the image. It is then assumed that the components near the frequency of the structure are scatter free and used to infer the primary component.

The remainder of the techniques reviewed involve the use of experimentation or simulators to estimate a Point Spread Function (PSF) (for example [62, 64, 66–70]). Then the PSFs are combined with Fourier deconvolution to predict and remove the scattering components. These techniques have a number of issues:

1. a small set of PSFs are assumed to be valid for the entire entity,

2. the PSFs are often assumed to be isotropic and stationary or isotropic and piece-wise stationary,

3. the PSFs are only dependent on the thickness of an entity not its chemical composition,

4. whether an acceptable solution has been found is not well tested,

5. the entity composition is not inferred and

6. the discrepancy between the simulator and the real world is not considered.

The assumption of an isotropic point spread function is a simplification to circumvent the computational challenges involved in scatter estimation. To deduce an isotropic PSF, a relatively simple simulation is required. Furthermore, the computational advantages of Fourier space convolution can be utilised. The approaches pass logical initial tests on simple entities but ignore the complex dependencies present and, therefore, it is only appropriate in a few constrained cases. These cases are not general enough to be relevant to the medical regime and, as such, these methodologies have not made it into common practice. [62] and [71] accept the existence of this problem and propose an adjustment to their PSFs to resolve the issue for mammography. Although an improvement, the adjustments are not sufficient to provide a satisfactory solution in the general case.

Another consideration is that the scattered photons blocked by the detector contain information about the entity's composition. Scatter may be used to resolve a fundamental degeneracy in X-ray physics, namely an observation of a thin strongly-attenuating material is identical to a thick weakly-attenuating material. Although true for the direct beam, this is not the case when scatter is considered. The scatter removal methods reviewed do not consider the extra information that can be gained by using scatter to make inferences about the entity's composition. If the information from scatter can be used to infer material and thickness, a thickness and material dependent scatter correction may be performed. Furthermore, this capability would open the door to making inferences about an entity's composition on standard digital radiography equipment, removing the need for referral for a multi-spectral examination.

In this chapter, a hierarchical-superposition emulator is used to resolve more uncertainty in system behaviour. In the following chapter, an inverse problem solving approach will be used to infer the entity's composition and improve scatter prediction when the entity's composition is uncertain. In Chapter 6, the difference between the model of X-ray examination and reality will be considered. Together this provides a method for scatter

removal and material inference which is better than an ASG and feasible for a wide variety of entities.

In the following section, two methods for modelling scatter will be set out. The first is analogous to current state-of-the-art in industry and the second is the full hierarchical-superposition strategy proposed in this thesis. For each case, i) the model is proposed and the assumptions are stated ii) the strategy for building the emulators is presented and iii) the method is validated using simulated X-ray images.

4.3.2 The Simulator: GEANT4

The simulator used in this thesis is GEANT4 (G4). G4 is a stochastic simulation toolkit which models the passage of particles through matter [113]. It is a flexible model which may be used for many purposes but this particular build was designed to mimic a medical X-ray examination system. A simulation produces a number of photons (events) at the source and tracks their passage until they are absorbed. The substance through which a photon is travelling dictates the probability of an interaction occurring. If a photon is absorbed by a detector, the energy deposited in the detector is recorded in the simulator output. G4 has many inputs which adjust the experimental set up and the entity examined. The interest of this chapter is to be able to build an emulator which resolves sufficient uncertainty to be able to improve on state-of-the-art scatter estimation. All the experimental inputs, apart from the beam divergence, were therefore fixed at an appropriate regime for medical imaging. When a simulation is run, a direction is sampled and the photon is fired in that direction; the beam divergence controls the set of directions from which the samples are made.

One reason an emulator is an essential tool in this application is G4 is expensive to evaluate. The most common simulation used in this thesis is one billion photons (events) which are fired in the same direction at a relatively simple entity. This is referred to as a pencil beam simulation and takes around ten minutes to run on a high-quality workstation. However, the speed is highly dependent on the complexity of the entity with more complex samples taking much longer to simulate. Furthermore, if the beam divergence is increased, the number of events required to get an image linearly increases with the number of pixels illuminated by the divergent beam. In summary, pencil beam simulations are expensive and many are required to model an X-ray image. Divergent beam simulations are very expensive and the cost of both pencil beam and divergent simulations scale badly with

sample complexity. This implies that using the simulator directly to perform inferences would not be feasible and an emulator would enable analysis in a useful time frame.

In the following two chapters, the simulator G4 is described by

$$\mathcal{F}(S^{G4}, M)$$

where $\mathcal{F}$ is the mathematical model encoded in G4 and S^{G4} is a three dimensional model of an entity with associated chemical composition encoded with a label, for example aluminium. The output $\mathcal{F}(S^{G4}, M)$ is the intensity recorded at every pixel on a (105×88) detector so $\mathcal{F}(S^{G4}, M)$ is a set of locations on a field which may be indexed spatially. M is the mode describing beam divergence where $M = 0$ is a five degree divergent beam and $M = (i, j)$ is a pencil beam directed at the centre of pixel (i, j).

Entity Dependent Parameters

In this build of G4, entity geometries are defined by a limited set of relatively simple shapes. The entity material is defined by a density and a material label which corresponds to a database of elements and alloys. The aim of the study is to be able to predict scatter for human anatomy, for example a Neck Of Femur (NOF). These entities are much more complex than can be included in this G4 build. In other words, the space to which S^{G4} belongs is a subset of the space to which S belongs. A relationship must be defined between an entity composition in the real world and the entities which can be simulated. The parameters will be referred to as entity dependent parameters and are notated x_S. The relationship between the entity parameters and an entity is described by $x_S = \mathcal{D}(S)$ where $x_S(i, j)$ is a two-vector containing the average thickness and material photons traverse in a square pyramid (SP) with its base on a pixel and point at the source. Here average material means the alloy of two materials as defined by G4. For example, the alloy could be tissue and bone and the average material is the proportion of bone, where 0 is 100% bone and 1 is 0% bone.

Figure 4.4 shows a diagram of the G4 set up. The SPs for each pixel are shown by the blue lines. The region of S that is averaged is illustrated by the blue truncated SP. $\mathcal{D}$ takes the average thickness and average material of the equivalent blue region for every SP.

The entity dependent parameters are written

$$\mathcal{D}(S) = x_S = \{\mu(i, j), \eta(i, j)\}_{i=1:\kappa, j=1:\theta}. \tag{4.3.1}$$

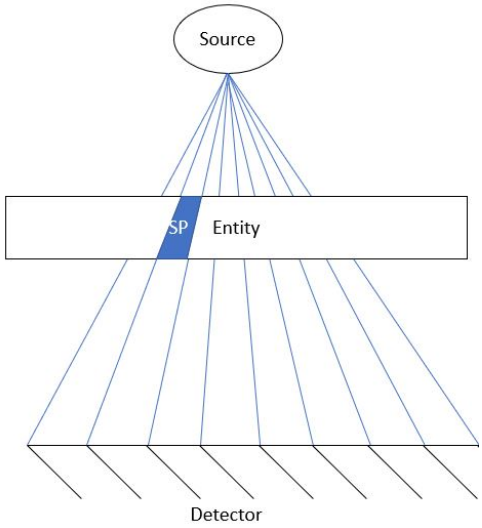

Figure 4.4: A Diagram of GEANT4. The blue lines show the area seen by each pixel. The blue box labelled SP shows a truncated square pyramid which is averaged in the discretisation process.

Here $\mu(i,j), \eta(i,j)$ is the average material and thickness a photon traverses in the SP with base on pixel (i,j). The definition of x_S was chosen for a number of reasons: i) it was the expected minimum number of parameters necessary to describe scatter accurately enough to provide image quality benefits, ii) it was the minimum number of parameters required so that diagnostic measures can be inferred and iii) a full three dimensional model was beyond the scope of this study. The assumption is that if

$$\begin{gathered} \mathcal{D}(S) = \mathcal{D}(S') \text{ then} \\ \mathcal{F}(S, M=0) = \mathcal{F}(S', M=0) + \epsilon \end{gathered} \tag{4.3.2}$$

where ϵ is an expectation zero random quantity the variance of which is small enough to have an insignificant impact on the inferences performed. $\epsilon = 0$ is true in G4 for the direct beam component. To see why this is the case consider equation 3.5.1. If the entity is made of two stacked cuboids with material μ_1 and μ_2, and thickness η_1 and η_2, the transmission is given by $\mathcal{T} = \exp\left(-(\mu_1(E)\eta_1 + \mu_2(E)\eta_2)\right)$. If the ordering of the materials is reversed, the transmission does not change as $\mu_1(E)\eta_1 + \mu_2(E)\eta_2 = \mu_2(E)\eta_2 + \mu_1(E)\eta_1$. This is known as the mixtures rule and holds in G4.

For the scatter component, it is non zero so the entities used for training the emulator are designed to be as similar as possible to the expected entities of interest. For example, if S is a patient's arm, the training entities are designed as a sandwich of tissue and bone. A sensitivity analysis was performed to validate that assuming $\epsilon = 0$ would not have a large impact on the inferences performed. This involved performing a number of simulations

of an image of a sandwich of two materials as shown in Figure 4.5. The total thickness was fixed to 7cm and the position of the central material varied from top to bottom. The percentage error between a simulation with an off centre central block and a simulation with a centred central block was below 5%. This was considered an acceptable error for the goals of this study. A more complete analysis with variable thickness and material would be an informative extension. If the error becomes unacceptable, the position of the central block could be added as an additional input. This would increase the dimension which would make emulation and inverse problem solving (see Chapter 5) more difficult, but it would relax the assumptions made.

To be able to evaluate G4 for arbitrary x_S, a separate object the shape of a truncated SP would be required for every pixel. Each truncated SP would also have to have a variable height and variable material present in it. This would be extremely challenging to produce in G4 for a number of reasons: i) the simulations would be dramatically more expensive, ii) the chance of a coding error would be high and iii) the precision of placement is too low. [1] Therefore, given that i) it would be too expensive to simulate, ii) the version of G4 available could not simulate it without significant further development and iii) the low precision would cause errors, it was not possible to simulate entities which are as complex as those of interest in the real world. Given this, the subsequent analysis is viewed as making inferences about a more complicated hypothetical simulator which has a higher degree of freedom in entity composition. The hypothetical simulators are defined as

$$\mathcal{F}'(x_S, M = 1) = Y^{Dr} \text{ and} \tag{4.3.3}$$

$$\mathcal{F}'(x_S, M = 2) = Y^{Sc}. \tag{4.3.4}$$

Here the prime notation is to differentiate the available simulator and the hypothetical simulators about which inferences will be made. $\mathcal{F}'$ may be evaluated for simple entities using a divergent beam simulation. These simulations require bespoke code for the entity geometry and take weeks to run on a standard workstation. Emulators for $\mathcal{F}'$ are validated using a couple of simulations of $\mathcal{F}'$ later in this chapter.

The method is to model the hypothetical simulator as a superposition of pencil beam simulations of entities with geometries simple enough to simulate efficiently in G4. This

[1] In an preliminary study the entity was coded in such a way that a pencil beam was directed into the gap between constituent objects caused by lack of precision.

approach is analogous to the use of Point Spread Functions (PSF) and convolution to model imaging systems. The current state-of-the-art uses pencil beam simulations of a cuboid entity at a number of discrete thicknesses. The simulator output is then calculated using piece-wise convolution on a segmented image. An equivalent approach using the Bayesian emulation methodology will be proposed in Section 4.3.3, its inadequacy demonstrated and an extension to overcome these problems presented in Section 4.3.4.

4.3.3 Isotropic Scatter Emulation

In this section, an equivalent approach to the current state-of-the-art for predicting scatter will be developed. Current approaches use a piece-wise stationary isotropic point spread function so this section will build a hierarchical-superposition emulator with non-stationary isotropic assumptions and demonstrate its inadequacy for the purposes of this study using the simulations of $\mathcal{F}'$. The simulators about which inferences are made are

$$Y = \mathcal{F}'(x_S, M = 0) = \mathcal{F}'(x_S, M = 1) \oplus \mathcal{F}'(x_S, M = 2).$$

Emulators for $\mathcal{F}'(x_S, M = 1)$ and $\mathcal{F}'(x_S, M = 2)$ will be built separately and referred to as the direct and scatter emulator. It is assumed that $\mathcal{F}'(x_S, M = 1)$ is uncorrelated with $\mathcal{F}'(x_S, M = 2)$. In fact they are likely to be positively correlated. However, the relationship is complex and unlikely to have a large impact on the inferences performed unless there is a large amount of uncertainty in the simulators. Given that the number of training points used to learn about $\mathcal{F}'$ and, as discussed in Section 4.3.5, that the uncertainty judgements are impractical to calculate, the additional work required to add in correlation was not considered worth the cost-to-benefit.

The Direct Component Model

The simpler of the two components to emulate is the direct beam as the output of a simulation at pixel t – where the simulation is a pencil beam directed at location t – only depends on $x_S(t)$. This means it can be assumed that the direct beam component at pixel t is independent of the entity given $x_S(t)$ and the response of the detector is spatially invariant. Therefore, the active parameters are defined as $x_a(t) = (\mu(t), \eta(t))$. Given these assumptions, a location on the full simulator output may be written as

$$\mathcal{F}'_{(t)}(x_S, M = 1) = \mathcal{F}_{(t^c)}(\mathcal{G}(x_a(t)), M = t^c) + \epsilon. \tag{4.3.5}$$

Here i) $t^c = (\frac{\kappa}{2}, \frac{\theta}{2})$ is the centre of the detector, ii) κ and θ are the number of pixels on each side of the detector and iii) ϵ is an expectation zero random quantity which describes scatter back in from the surrounding area. ϵ was shown to be negligible using a number of pencil beam simulations with the geometry given in Figure 4.9. The composition of the entity at in the centre was kept constant and the surrounding material varied. This was repeated for a number of entities and the error was not differentiable from Poisson noise.

The primitive emulator for the direct beam is given by

$$\log(\mathcal{F}_{(t^c)}(\mathcal{G}(x_a(t)), M = t^c)) = \mathcal{M}^{Dr}(x_a(t)) \oplus \mathcal{W}^{Dr}(x_a(t)). \tag{4.3.6}$$

Here $\mathcal{G}(x_a(t))$ defines a cuboid entity covering the entire detector with thickness $\eta(t)$ and material $\mu(t)$. The log transformation is one often used in this work and it is motivated by the Beer-Lambert law given in equation 3.5.1. The log of transmission $\log(\mathcal{T}) = \mu\eta$ is linear for this model. As G4 is poly-chromatic, see equation 3.5.2, the logarithmic transform does not linearise the data but makes it almost quadratic. Therefore, a quadratic mean function is used with a correlated residual process.

Figure 4.5 shows a representation of the entity geometry defined by $\mathcal{G}(x_a(t))$. The entity is a sandwich of materials so $\mu(t)$ is the proportion of material one. Using this model, the Direct-Beam-Emulator may be produced from $\kappa\theta$ (one simulation for every pixel on the detector) evaluations of an emulator with two input dimensions and a scalar output.

The Scattering Component Model

The scattering component is more challenging as every SP affects every location on the detector. The non-stationary isotropic approach also assumes that the output of a pencil beam simulation directed at t is independent of x_S given $x_S(t)$. It is assumed that the simulator is equal to a summation of pencil beam simulations

$$\mathcal{F}'(x_S, M = 2) = \sum_{i=1}^{\kappa} \sum_{j=1}^{\theta} \mathcal{F}'(x_S, M = (i, j)) \tag{4.3.7}$$

where $\mathcal{F}'(x_S, M = (i, j))$ is a simulation of a pencil beam directed at pixel (i, j). It is then assumed that

$$\mathcal{F}'_{(t^{(i)})}(x_S, M = t^{(j)}) = \mathcal{F}'_{(t^{(k)})}(x'_S, M = t^{(l)}) + \epsilon \tag{4.3.8}$$

given $x_S(t^{(j)}) = x'_S(t^{(l)})$ and $|t^{(i)} - t^{(j)}| = |t^{(k)} - t^{(l)}|$. In words, the output of the simulator at pixel $t^{(i)}$ from a pencil beam simulation directed at pixel $t^{(j)}$ is equal to the output of

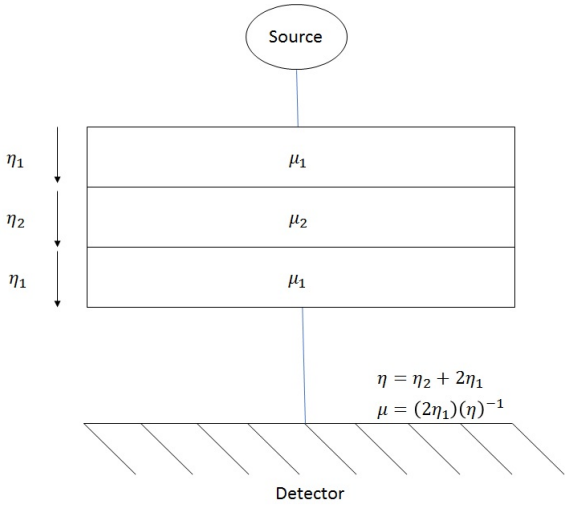

Figure 4.5: A diagram illustrating the entity set up for the isotropic $\mathcal{G}(\mu, \eta)$. The blue line represents a pencil beam fired from the source at the centre of a detector. The entity is made of a sandwich of two materials. μ_1 is material one, μ_2 is material two, η_1 is the thickness of material one and η_2 is the thickness of material two.

the simulator at pixel $t^{(k)}$ from a pencil beam simulation directed at pixel $t^{(l)}$ given that the entities' compositions at location $t^{(j)}$ and $t^{(l)}$ are the same. In this case, ϵ should not be ignored and a residual process should be added using the method outlined in Section 4.2.3. However, the aim of this section to is match the current state-of-the-art mathematical scatter removal methods to be able to demonstrate the necessity of the improvements presented in Section 4.3.4. Therefore, for the non-stationary isotropic method, it will be assumed that $\epsilon = 0$ and in the next section, when the method used throughout the rest of this thesis is presented, it will not. The relationship between $\mathcal{F}'$ and $\mathcal{F}$ is given by

$$\mathcal{F}'_{(t^{(i)})}(x_S, M = t^{(j)}) = \mathcal{F}_{(t^{(k)})}(\mathcal{G}(x_a(t^{(j)})), M = t^c) \text{ for } t^{(i)} \neq t^{(j)}$$

where $|t^{(i)} - t^{(j)}| = |t^c - t^{(k)}|$. In words, the output of the simulator at pixel $t^{(i)}$ from a pencil beam simulation directed at pixel $t^{(j)}$ is equal to the output of the simulator at location $t^{(k)}$ from a pencil beam simulation directed at the centre of the detector through a cuboid with thickness $\eta(t^{(j)})$ and alloy $\mu(t^{(j)})$. The primitive emulator is

$$\log(\bar{\mathcal{F}}_{(t)}(\mathcal{G}(x_a(t)), M = t^c)) = \mathcal{M}^{Sc}(x_a(t), |t - t^c|) \oplus \mathcal{W}^{Sc}(x_a(t), |t - t^c|) \tag{4.3.9}$$

where $\bar{\mathcal{F}}_{(t)}(\mathcal{G}(x_a(t)), M = t^c)$ is the mean of all pixels $r = |t - t^c|$ from t^c. The average is used because $\mathcal{G}$ produces radially symmetric entities and the simulator is stochastic, so taking the radial average resolves more uncertainty in the simulator's population mean.

This means that fewer photons are required to get a sufficiently small signal-to-noise-ratio and hence the simulator is more efficient. In this case, the adjusted expectation of $\mathcal{F}'(x_S, M = 2)$ is produced by $(\kappa\theta)^2$ evaluations of a scalar emulator with three input dimensions. Building an emulator with three input dimensions and one output dimension is easier than building an emulator with $\theta\kappa$ input dimensions and $\theta\kappa$ output dimensions. This is why these types of approaches have been employed for this problem.

Now that the model for the non-stationary-isotropic method has been proposed, the next steps are to build the direct and scatter emulators, validate those emulators and test the emulator of $\mathcal{F}'$ using divergent beam simulations.

Building The Direct Emulator

The mean function of the direct emulator, given in equation 4.3.6, is

$$\mathcal{M}^{Dr}(x_a(t)) = (1, x_a(t), x_a(t)^2)\beta^{Dr}$$

where β^{Dr} is a vector of regression parameters. The residual process correlation structure was specified using an exponential family correlation function, given in equation 2.2.8, with correlation length $l = (0.2, 3)$ and $p = 2$. The regression parameters and variance multiplier were fixed using the GLS estimators, see key-formula 2.2.1. The emulator was trained using 108 pencil beam simulations and validated using cross validation. The 108 simulations were chosen using a lattice design for which the points of interest were not near the edge of the set. The points of interest were the x_S associated with the entities used for divergent beam validation simulations shown in Figure 4.6. This model has been used on a number of data-sets for a range of thicknesses and alloys, a subset of which are presented in this thesis. The only parameter which sometimes requires alteration is the correlation length if the sparsity of the training-set changes.

Building the Scatter Emulator

The form of the mean function requires more care as the gradient in the $|t - t^c|$ dimension is steep enough that polynomials did not resolve enough non-stationary behaviour. The reason is that the logarithmic transform does not linearise the behaviour as well as the direct component. This is because there are three complex processes, two of which have an exponential component. The description in this paragraph is not used directly but illustrates why the scatter emulator was more challenging to fit. Consider a pencil beam

travelling through an entity, at a given location within the entity, photons are scattered. This can be thought of as another source producing photons. Now consider a location on the detector which r away from the scattering centre. The spectrum at the scattering centre is an application of the Beer-Lambert law, given in equation 3.5.2, on the incident spectrum $BL(I_0, x_S(t)) = Sp_1$. Sp_1 is then passed through a complex function $Sp_2 = KN(Sp_1, r)$ which dictates the spectrum produced travelling in the direction which would hit a pixel distance r from the scattering centre. Finally, this spectrum is again acted on by the Beer-Lambert law $BL(Sp_2, x_S(t))$. The fact this attenuation only depends on $x_S(t)$ is expressly the assumption made in the isotropic approach. Overall the scatter produced as a function of r looks like a double exponential.

Given these considerations, the form of the mean function for the scatter emulator, given in equation 4.3.9, is

$$\begin{aligned}\mathcal{M}^{Sc}(x_a(t), r) =&(1, x_{a(1)}(t), x_{a(2)}(t), x_{a(1)}(t)x_{a(2)}(t), x_{a(1)}(t)^2, x_{a(2)}(t)^2, \log(r+1), (r+1)^{-1}, r,\\ &\log(r+1)x_{a(1)}(t), (r+1)^{-1}x_{a(1)}(t), rx_{a(1)}(t),\\ &\log(r+1)x_{a(2)}(t), (r+1)^{-1}x_{a(2)}(t), rx_{a(2)}(t))\beta^{Sc}\end{aligned}$$

which accounted for the majority of the variation. β^{Sc} is a vector of regression parameters and r is the distance from the scattering centre. The residual process correlation structure was specified using an exponential family correlation function, given in equation 2.2.8, with $l = (0.2, 3, 10)$ and $p = 2$. The regression parameters and variance multiplier were fixed using the GLS estimators, see key-formula 2.2.1. An alternative approach would have been to perform two logarithmic transforms. However, as the simulator output could be less than one, this was not robust. Another choice was to fit an exponential to the log of the simulator output. This would mean the GLS approach could not be applied. As both these avenues were challenging, the mean function was chosen by testing a number of forms and deciding which one, in conjunction with the residual process, validated best. This involved graphical plots as well as the diagnostic measures set out in Section 2.2.2.

Isotropic Results

Given valid emulators for both direct and scattering components, beliefs about the hypothetical simulator can be adjusted. To test this method, divergent beam simulations of two entities were used: one has a simple geometry designed so that the isotropic assumption is appropriate and the second has a more complex geometry designed so that the isotropic assumption is inappropriate. Figure 4.6 shows the output of $\mathcal{F}'(x_S, M = 0)$

for the validation entities and diagrams of the entities' geometries.

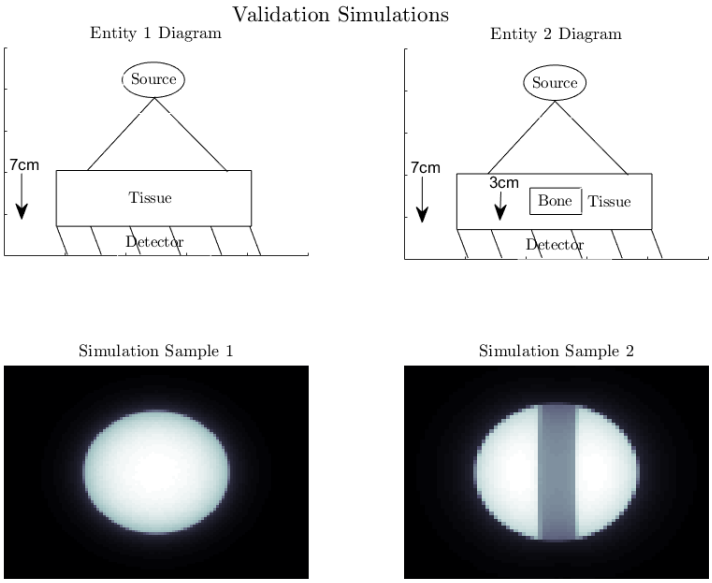

Figure 4.6: Top left: a diagram of a validation sample one. Top right: a diagram of a validation sample two. Bottom left: an image of simulator output for validation entity one. Bottom right: an image of simulator output for validation entity two.

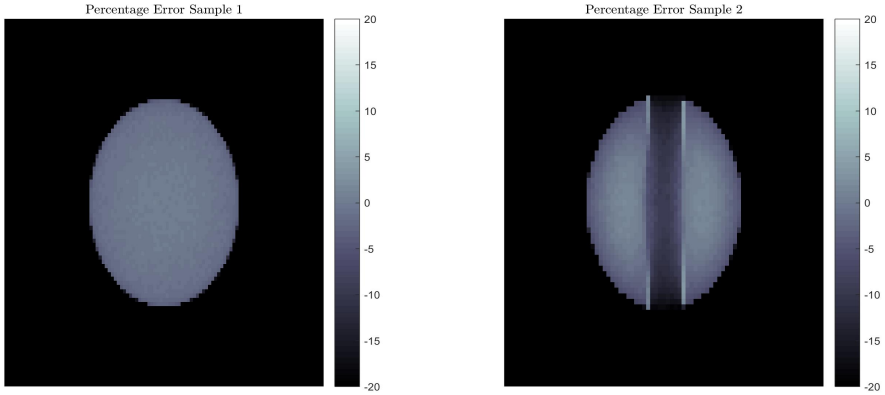

Figure 4.7: Left: percentage error between simulation and adjusted expectation of a simulation for validation entity one. Right: percentage error between simulation and adjusted expectation of a simulation for validation entity two.

Figure 4.7 shows the percentage error between simulator $\mathcal{F}'$ and the adjusted expectation of $\mathcal{F}'$ for the validation simulations. The left image shows the cuboid of tissue equivalent material where the isotropic assumption is valid. The error is within 1% which

augers well for resolving sufficient uncertainty to return useful diagnostic quantities and improve image quality. The right image shows the same block of tissue with the central portion removed and replaced with bone equivalent material. The error has spiked to within 20% and contains clear entity dependence.

The advantage of the isotropic assumption is that it is simple and looks promising given logical initial tests on simple entities. However, it is not sufficient for slightly more complex entities. Given that the aim is to perform this analysis on human anatomy and return useful diagnostic measurements, this level of emulator error is unacceptable. The cause of the error is that the assumptions in the emulation process were too strong and therefore must be relaxed.

4.3.4 Non-Isotropic Scatter Estimation

The isotopic technique for scatter estimation does not produce an emulator with an error consistently small enough to achieve the goals of this study. In this section, the causes of the error will be illuminated, a hierarchical-superposition strategy proposed and the improvement quantified. To illustrate the issue, consider a pencil beam travelling through an entity. At a given location scattered photons are produced, this location will be referred to as a scattering centre. The scattering centre can be considered as a new source which produces photons of a given spectrum. The isotropic method models the spectrum that is produced at a scattering centre well, however, it assumes that the radial material is the same as the centre. Therefore, the attenuation of the scattered photons is not well modelled when the composition differs from that of the scattering centre. This attenuation of scattered photons will be referred to as radial attenuation.

Figure 4.8 illustrates the radial attenuation problem for validation entity two. It shows the error between the emulator and the simulator for a pencil beam simulation directed at $(i, j) = (44, 42)$. In the region where the entity has the same composition as $(\eta(44, 42), \mu(44, 42))$, the error is small. However, if the radial composition is different to the composition at the scattering centre, the error increases. This can be seen by an increasing error moving left to right across the right hand image in Figure 4.8. Essentially the isotropic assumption induces too large an error. In order to relax this assumption, a new emulation strategy will be defined which parameterises the remaining dependence on x_S into a concise form. Then a new entity geometry will be proposed which allows the majority of the variation to be resolved and the emulator uncertainty to be reduced to a

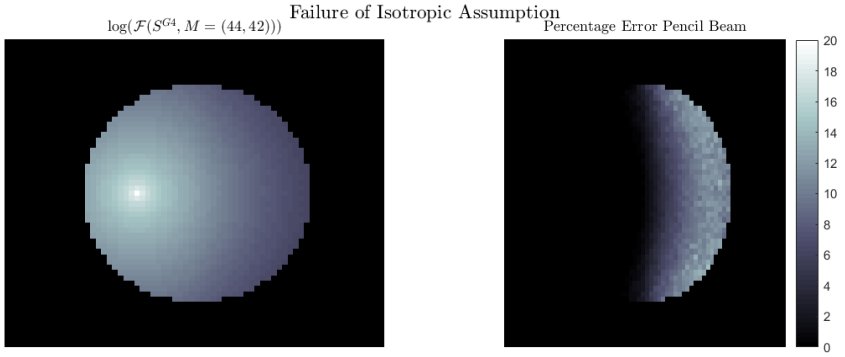

Figure 4.8: Left: a pencil beam simulation directed at pixel (44 , 42) with entity 2 configuration as shown in Figure 4.7. Right: the percentage error between the pencil beam simulation and the adjusted expectation of the simulator given the isotropic assumption.

useful level. Finally, a residual process will be added to account for the error induced by the assumption.

The Non-isotropic Model

The model for the direct beam simulator was considered sufficient and did not require alteration. The scatter component on the other hand requires a number of improvements. To begin the construction of the non-isotropic hierarchical-superposition emulator of G4, it is assumed that 4.3.7 holds. For this approach, the expected intensity deposited on pixel $t^{(i)}$ from a pencil beam directed at pixel $t^{(j)}$ is assumed to be the same as a pencil beam simulation with any entity which has scattering centre composition $x_S(t^{(j)})$ and radial composition $u^{(i,j)} = (\mu_t^{(i,j)}, \eta_t^{(i,j)})$. $(\mu_t^{(i,j)}, \eta_t^{(i,j)})$ is the average projected material and thickness between $t^{(i)}$ and $t^{(j)}$. In other words, given two entities have the same composition for the scattering centre and the same average composition on the path between scattering centre and a pixel on the detector, the intensity at that pixel will be approximately equal for both simulations. More formally, a function $\mathcal{TR}$ is designed so that it is reasonable to assert that if

$$\mathcal{TR}(x_S, t^{(i)}, t^{(j)}) = \mathcal{TR}(x'_S, t^{(k)}, t^{(l)}), \text{ then}$$
$$\mathcal{F}'_{(t^{(i)})}(x_S, M = t^{(j)}) = \mathcal{F}'_{(t^{(k)})}(x'_S, M = t^{(l)}) + \epsilon$$

where ϵ has expectation zero. The function in this case is $\mathcal{TR}(x_S, t^{(i)}, (t^{(j)}) = (x_a(t^{(j)})), u^{(i,j)})$. In words, the output of the simulator at pixel $t^{(i)}$ from a pencil beam simulation directed at pixel $t^{(j)}$ is equal to the output of the simulator at pixel $t^{(k)}$ from a pencil beam sim-

ulation directed at pixel $t^{(l)}$ given that the entities' compositions at location $t^{(j)}$ and $t^{(l)}$ are the same and the average projected radial material is also the same.

Although this assumption is better than assuming the radial material is the same as the scattering centre, it is still only approximately true. Therefore, a residual process is added to take ϵ into account. In order to create a simulation with a scattering centre composition defined by $x_a(t^{(j)})$ and radial composition $u^{(i,j)}$, a new design of entity geometry $\mathcal{G}$ was required. Figure 4.9 depicts the entity geometry designed to be able to get a simulation informative for $\mathcal{F}'$. The central core and the surrounding material are both a sandwich of two materials where each material has a given thickness. This enables for any $(x_a(t), u)$, a simulation to be produced with scattering centre $x_a(t)$ and radial composition u. $\mathcal{F}'$ is therefore related to G4 by

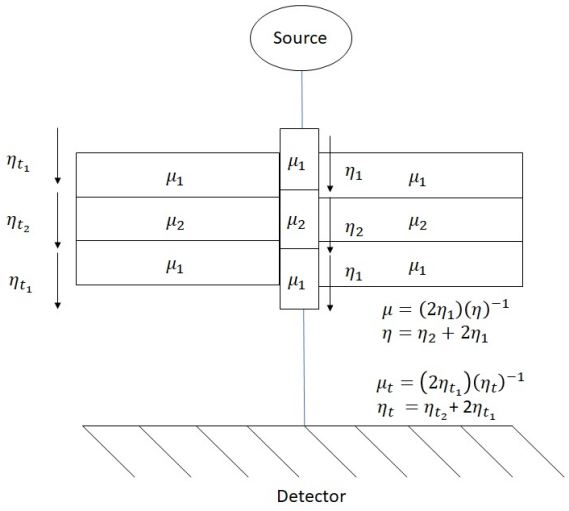

Figure 4.9: Diagram illustrating the entity design $\mathcal{G}(x_a(t), u)$ which is used for the non-isotropic emulation approach.

$$\mathcal{F}'_{(t^{(i)})}(x_S, M = t^{(j)}) = \mathcal{F}_{(t^{(k)})}(\mathcal{G}(x_a(t^{(j)}), u^{(i,j)}), M = t^c) \oplus \mathcal{W}^a(x_a(t^{(j)}), u^{(i,j)}, |t^{(i)} - t^{(j)}|) \tag{4.3.10}$$

where $\mathcal{G}$ defines the entity geometry shown in Figure 4.9 and $|t^{(i)} - t^{(j)}| = |t^{(k)} - t^c|$. In words, the output of the simulator at pixel $t^{(i)}$ from a pencil beam simulation directed at pixel $t^{(j)}$ is equal to the output of the simulator at pixel $t^{(k)}$ from a pencil beam simulation directed at the centre of the detector through a sample with core material $x_a(t^{(j)})$ and radial material $u^{(i,j)}$ plus an expectation zero residual process. When $u^{(i,j)} = x_a(t^{(j)})$, the non-isotropic and isotropic versions are equal, therefore, the isotropic approach can

be seen as nested in the non-isotropic approach. Then the primitive emulator is given by

$$\log(\bar{\mathcal{F}}_{(t)}(\mathcal{G}(x_a(t), u), M = t^c)) = \mathcal{M}^{Sc}(x_a(t), u, |t - t^c|) \oplus \mathcal{W}^{Sc}(x_a(t), u, |t - t^c|)$$

where $\bar{\mathcal{F}}_{(t)}(\mathcal{G}(x_a(t), u), M = t^c)$ is the average of all pixels $|t - t^c|$ from t^c. The next step is to be able to calculate the average radial entity thickness and material $u^{(i,j)}$. This is found by a heuristic ray tracing algorithm, notated by

$$\mathcal{RT}(x_S, t^{(i)}, t^{(j)}) = u^{(i,j)}.$$

The pseudo ray-trace algorithm is given by Algorithm 1.

Algorithm 1 Pseudo Ray-trace.

Inputs: x_S, $t^{(i)}$, $t^{(j)}$, $\tilde{t}$

Outputs: $u^{(i,j)}$

1: Calculate the first time a pencil beam directed at $t^{(i)}$ enters the entity and define the scattering centre as $\tilde{t}$cm into the entity.

2: Calculate the pixel below the location where a pencil beam directed from the scattering centre to pixel $t^{(j)}$ exits the entity, notated $t^{(k)}$.

3: Calculate the set of pixels which intersect a line between $t^{(k)}$ and $t^{(i)}$, notated T^*.

4: Average $x_S(T^*)$ to get the projected radial material and thickness $u^{(i,j)}$.

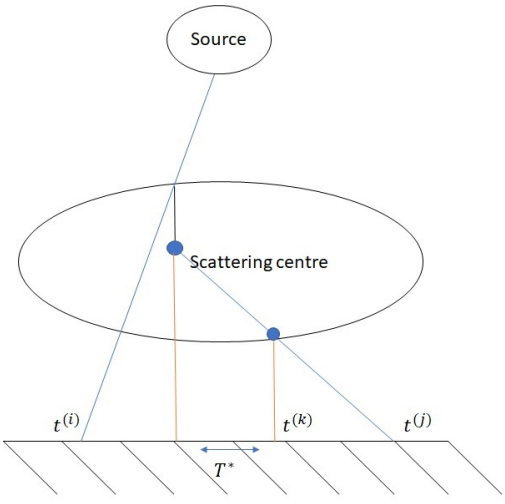

Figure 4.10: Diagram illustrating how T^* is calculated in the ray trace algorithm.

Figure 4.10 shows a diagram of how T^* is calculated in the ray trace algorithm. For the examples presented in this section, $\tilde{t}$ was set at 2.6cm. This was chosen because

simulations showed that the most scatter is produced 2.6cm into a 7cm sample of PMMA. If the entity is directly on the detector, the algorithm output is invariant to $\tilde{t}$. As, in this work, the entities are always placed on the detector or very close to it, $\tilde{t}$ does not have a large impact in this thesis. For applications like Computed Tomography (CT), the effect could be much larger as entities are placed centrally between source and detector.

This ray tracing procedure was motivated by the idea that if the material and thickness on every path between scattering centre and a location on the detector was known, the intensity at the detector could be modelled. There were three problems with performing the three dimensional ray trace. The first is that it is prohibitively expensive. The second is that it can only be done if a three dimensional representation of the entity is available and, in the real world application, no known representation would be. Finally, the output is two dimensional so making inferences for a three dimensional structure from a two dimension projection would be hard to solve. The approach here was a cheaper alternative which still gave additional parameters which are correlated with the material composition along a ray. Although heuristic, it has been shown that these additional parameters enable a larger amount of simulator uncertainty to be resolved and the error between the emulator and $\mathcal{F}'$ to be reduced. This has been shown to outperform current methods on a number of simulations where there is a known truth and results in the real world look plausible (see Chapter 6).

The set of reduced simulator inputs and outputs is given by

$$\mathcal{D}'(x_S) = \{x_a(t^{(i)}), u^{(i,j)}, r^{(i,j)} = |t^{(i)} - t^{(j)}|\}_{i=1:\kappa\theta, j=1:\kappa\theta}.$$

The calculation of the adjusted expectation of $\mathcal{F}'(x_S, M = 0)$ requires i) $(\kappa\theta)^2$ evaluations of $\mathcal{RT}$, ii) the calculation of $(\kappa\theta)^2$ distances, iii) $(\kappa\theta)^2$ evaluations of a five dimensional emulator and iv) $(\kappa\theta)$ evaluations of a two dimensional emulator.

Now calculating the adjusted expectation of $\mathcal{F}'$, which has input dimension $2\kappa\theta$ and output dimension $\kappa\theta$, may be performed with $(\kappa\theta)^2$ evaluations of scalar emulators with five input dimensions and $\kappa\theta$ evaluations of an emulator with two input dimensions. The training of the emulator with five input dimensions is challenging and required a new emulation strategy.

Building The Non-Isotropic Emulator

A large (over 10000) number of training points were required so that the error between $\mathcal{F}'$ and its adjusted expectation was small enough for the subsequent goals of the study to be

met. Due to the size of this training-set and the number of emulator evaluations required, using a conventional emulator would still be too expensive for an *real-time* algorithm so the look-up-table approach was used. Instead of needing an efficient emulator, the requirement is a reasonably efficient method to create a look-up-table which utilises all available simulations. How the method will be eventually deployed in hospitals is still being decided. Depending on the stability of the equipment, a look-up-table may only need to be built once a week or maybe multiple times a day. If the table takes several hours to create, clinical practice might be impeded and the method therefore unusable. Therefore, an efficient method of building look-up-tables will facilitate the adoption of the methods in industry. The process is as follows: i) design the look-up-table utilising the available memory, ii) design an emulation strategy to calculate the look-up-table, iii) run simulations iv) build and validate the emulator and v) evaluate the emulator to create the look-up-table. The look-up-table is then the *de-facto* emulator for the online version of the algorithm.

In this example, the look-up-table was designed using 100^4 points in a lattice. 100^4 was chosen as it was enough points to fill the space and small enough to store in virtual memory on a workstation. The lattice was chosen so that all values in $\mathcal{D}'(x_S)$, for both validation entities shown in Figure 4.6, were in the convex hull and not too near the edge. The simulations were chosen to span the convex hull of the look-up-table, also in a lattice. This means that 5184 simulations were used to build the scatter emulator. This strategy is used for building look-up-tables throughout the rest of the thesis.

Given this design, the next step is to process the data into a form which could be emulated. The emulator training-set is $\{X, F\}$ where

$$X = \{(x_a(t^{(i)}), u^{(i)})\}_{i=1:5184} \text{ and}$$
$$F = \{\mathcal{F}_{(t^{(j)})}(\mathcal{G}(x_a^{(i)}, u^{(i)}), M = t^c)\}_{i=1:5184, j=1:\kappa\theta}$$

where $\kappa\theta = 8976$ Given that i) the non-isotropic emulator only depends on the distance between pixels and ii) the population mean of $\mathcal{F}_{(t^{(j)})}(\mathcal{G}(x_a^{(i)}, u^{(i)}), M = t^c)$ is symmetric, the simulator output is radially collapsed about the central pixel t^c. Therefore, the emulator training-set is reformulated as

$$X = \{x_a(t^{(i)}), u^{(i)}, r^{(j)}\}_{i=1:5184, j=1:d} \text{ and}$$
$$F = \{\bar{\mathcal{F}}_{(r^{(j)})}(\mathcal{G}(x_a(t^{(i)}), u^{(i)}), M = t^c)\}_{i=1:5184, j=1:d}$$

where $d = 50$ is the number of distances $r^{(j)} = |t^{(j)} - t^c|$ on a radius and $\bar{\mathcal{F}}_{(r)}(\cdot, M = t^c)$ is the average of all pixels r from t^c. Given this simulation design strategy and the radial collapse, the number of points available for training is 259200. To circumvent memory bottlenecks in the emulation procedure, a two stage strategy was utilised: first emulate over r for each $(x_a(t^{(i)}), u^{(i)})$ and calculate the adjusted expectation for all unique $r^{(j)}$ in the look-up-table design, then emulate over $(x_a(t), u)$. The algorithm for building the look-up-tables is given by Algorithm 2. This algorithm is used for many examples

Algorithm 2 Build Look-up-tables.

Inputs: X, F, $X^{\mathcal{LU}}$, a, b, l_1, l_2

Outputs: $\mathcal{LU}$

1: **begin:** Emulate over r.
2: **for** i in 1:$\mathbb{N}$ **do**
3: Using l_1 , a , b build an emulator of $\log(\bar{\mathcal{F}}_{(r)}(\mathcal{G}(x_a(t^{(i)}), u^{(i)}), M = t^c))$
4: **for** j in 1:$\tilde{d}$ **do**
5: Calculate $\mathbb{E}_{F^{(i)}}[M(\log(\bar{\mathcal{F}}_{(r^{(j)})}(\mathcal{G}(x_a(t^{(i)}), u^{(i)}), M = t^c)))]$
6: **end for**
7: **end for**
8: **end**
9: Using l_2, build a multiple output emulator of

$$\left(\mathbb{E}_{F^{(i)}}[M(\log(\mathcal{F}_{(r^{(j)})}(\mathcal{G}(x_a(t^{(i)}), u^{(i)}, M = t^c)]\right)_{j=1:\tilde{d}}$$

over $(x_a(t^{(i)}), u^{(i)})$.
10: **for** i in 1:nrow($X^{\mathcal{LU}}$) **do**
11: Calculate the adjusted expectation of $\left(\mathbb{E}_{F^{(i)}}[M(\log(\mathcal{F}_{(r^{(j)})}(\mathcal{G}(x_a(t^{(i)}), u^{(i)}, M = t^c)]\right)_{j=1:\tilde{d}}$
12: **end for**

presented in the rest of this thesis. The form of the emulators will now be discussed and the choices of parameters for the examples in this chapter stated. The emulator of $\log(\bar{\mathcal{F}}_{(r)}(\mathcal{G}(x_a(t^{(i)}), u^{(i)}), M = t^c))$ is formulated as a stochastic emulator of the form.

$$\begin{aligned}\log(\bar{\mathcal{F}}_{(r)}(\mathcal{G}(x_a(t^{(i)}), u^{(i)}), M = t^c)) = M(\log(\bar{\mathcal{F}}_{(r)}(\mathcal{G}(x_a(t^{(i)}), u^{(i)}), M = t^c))) \\ \oplus R^{(j)}(\log(\bar{\mathcal{F}}_{(r)}(\mathcal{G}(x_a(t^{(i)}), u^{(i)}), M = t^c))).\end{aligned}$$

The population mean emulator was specified using

$$M(\log(\bar{\mathcal{F}}_{(r)}(\mathcal{G}(x_a(t^{(i)}), u^{(i)}), M = t^c))) = \left(1, r, r^2, \frac{1}{r+1}\right) \beta \oplus \mathcal{W}_M(r). \tag{4.3.11}$$

The population variance of the emulator was specified as

$$\frac{\mathbb{V}ar[R^{(j)}(\log(\bar{\mathcal{F}}_{(r)}(\mathcal{G}(x_a(t^{(i)}), u^{(i)}), M = t^c)))]}{\mathbb{V}ar\left[\mathcal{W}_M(r)\right]} = \frac{a}{\bar{\mathcal{F}}_{(r)}(\mathcal{G}(x_a(t^{(i)}), u^{(i)}), M = t^c)} + b. \tag{4.3.12}$$

This form was chosen as the simulator output is an integral over Poisson distributions. Therefore, the population mean and variance of the simulator can be assumed to be proportional to one another and the distribution can be well modelled as Gaussian. Then the log of the simulator output will have variance proportional to

$$\left(M(\log(\bar{\mathcal{F}}_{(r)}(\mathcal{G}(x_a(t^{(i)}), u^{(i)}), M = t^c)))\right)^{-1}.$$

Given the radial collapse and that the argument above is only approximately true, the form for direct specification was

$$a\left(M(\log(\bar{\mathcal{F}}_{(r)}(\mathcal{G}(x_a(t^{(i)}), u^{(i)}), M = t^c)))\right)^{-1} + b$$

where a , b are constants which were specified using validation procedures and graphical analysis. For the example in this section, $a = 0.0001$ and $b = 0.001$. Originally a full Bayes linear stochastic emulator (see Appendix A.6) was used to build the emulators. However, the direct specification was found to work as well in terms of percentage error and well enough for uncertainty estimation. This form has been used to build many look-up tables of many data-sets so the method is well validated. Issues occur when the entity composition attenuates too many photons so that the tails contain zeros. This causes problems with the log transform and stochastic component, but it can easily be fixed by simulating more events.

The regression parameters and variance multipliers were fixed to their GLS estimates using key-formula 2.2.1. The residual process correlation structure was specified using equation 2.2.8 with correlation length $l_1 = 6$ and $p = 2$. The approximate adjusted expectations of the components in equation 4.3.11 are then calculated for the $\tilde{d}$ unique distances in the look-up-table design.

The emulator of $\left(\mathbb{E}_{F^{(i)}}[M(\log(\mathcal{F}_{(r^{(j)})}(\mathcal{G}(x_a(t^{(i)}), u^{(i)}, M = t^c)]\right)_{j=1:\tilde{d}}$ is specified as

$$\left(1, x_a(t^{(i)}), u^{(i)}, x_a(t^{(i)})u^{(i)}, x_a(t^{(i)})^2, u^{(i)\,2}\right) \beta \oplus \mathcal{W}(x_a(t^{(i)}), u^{(i)}). \tag{4.3.13}$$

Here $F^{(i)}$ notates the subset of F such that $(x_a(t), u) = (x_a(t^{(i)}), u^{(i)})$. The residual process covariance structure was specified as separable (equation 2.2.7) with correlation function specified using equation 2.2.8 with correlation length $l_2 = [0.2, 1, 0.2, 1]$ and $p = 2$. GLS estimates, key-formula 2.2.1, were used to compute the approximate adjusted beliefs.

This strategy breaks up the emulator into 5184 scalar heteroscedastic emulators, equation 4.3.11, with 50 training points and one multiple output emulator with 5184 training points.

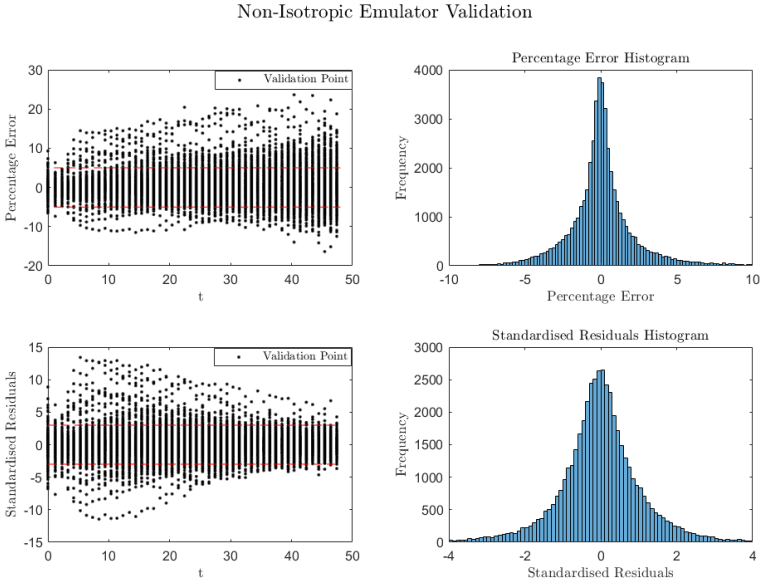

Figure 4.11: Top left: percentage error for every point in the validation set. Top right: percentage error histogram for every point in the validation set. Bottom left: standardised errors for every point in the validation set. Bottom right: histogram of standardised errors for every point in the validation set.

Figure 4.11 shows example validation plots for the primitive emulator. The percentage error is reasonably low with the majority of points falling within 5%. There is clear dependence with t (trumpet shape) which is as expected given the heteroscedasticity. The standardised residuals remove the majority of the trumpet shape. However, the uncertainty is generally slightly overestimated apart from closest to the scattering centre where it is underestimated. This is because the mean function does not resolve all non-stationary behaviour and the heteroscedastic model has a number of heuristic approximations.

This method is used to build many look-up tables in the next two chapters. The regression component is fixed in all and has been used for many applications of this method, both in simulation and the real world. The parameters l_1, l_2, a and b change

depending on the emulator training points. These are manually changed by the user if validation plots or diagnostics fail.

For the non-stationary-isotropic approach, a direct emulation strategy was sufficient as the dimension is small enough that a reasonable number of training points well represented the space. In the non-isotropic case, the set of simulations available would have to be significantly sub-sampled to be able to perform the calculations. The direct approach was attempted with 10000 carefully chosen points. The error between the simulator and emulator was slightly higher for validation points and the time to create the full look-up table took hours rather than minutes.

The summary of the non-stationary-non-isotropic approach to predicting scatter so far is separated into two algorithms. The first is building the emulators summarised in Algorithm 3. Then for prediction at a new location, the method is performed using Algorithm 4.

Algorithm 3 Build Emulators and Look-up-tables.

Inputs: X, $X^{\mathcal{LU}}$, a, b, l, l_1, l_2

Outputs: $\mathcal{LU}$, Direct-Beam-Emulator.

1: **begin:** Evaluate the simulator for all points in simulation design X to calculate F.
2: **for** i in 1:nrow(X) **do**
3: Evaluate $\mathcal{F}(\mathcal{G}(x_a^{(i)}(t), u^{(i)}), M = t^c)$.
4: Radially collapse the simulator output to calculate the i^{th} row of F.
5: **end for**
6: **end**
7: Using the central pixel $\mathcal{F}_{(t^c)}(\mathcal{G}(x_a^{(i)}(t), u^{(i)}), M = t^c)$ and correlation length l, build and validate the Direct-Beam-Emulator.
8: Using X , F, a , b , l_1, l_2 and $X^{\mathcal{LU}}$, build and validate the look-up-table using Algorithm 2.

It is now time to return to the example used to illustrate the inadequacy of the non-stationary-isotropic approach. The entity used in this divergent beam simulation is given in Figure 4.6 right. It is interesting to validate whether this algorithm, without the residual process $\mathcal{W}^a$, performs better than the isotropic case for the entity where the isotropic assumptions were invalid. Figure 4.12 and 4.13 show comparisons between emulators with the isotropic and non-isotropic assumptions. Figure 4.12 shows the percentage error for a pencil beam simulation for both techniques. The non-isotropic approach has a lower

Algorithm 4 Prediction for $\mathcal{F}'$ with Look-up-table.

Inputs: $\mathcal{LU}$, S', Direct-Beam-Emulator

Outputs: $\mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}'(x'_S, M=0)]$

begin: Calculate discretisation of S' $\mathcal{D}(S') = x'_S$.
Calculate $\mathcal{D}'(x'_S)$.
for i in 1:$\theta\kappa$ **do**
    for j in 1:$\theta\kappa$ **do**
        Set $x_a^{(j)}(t^{(i)}) = x_S^{(i)}$.
        Calculate radial distance $r^{(i,j)} = |t^{(i)} - t^{(j)}|$.
        Calculate $u^{(i,j)}$ using $t^{(i)}$, $t^{(j)}$, x'_S and Algorithm 1.
    end for
end for
end
begin: Evaluate look-up-table for all elements in $\mathcal{D}'(x'_S)$.
for i in 1:$(\theta\kappa)^2$ **do**
    Evaluate $\mathcal{LU}(x_a^{(i)}, u^{(i)}, r^{(i)})$
end for
end
Combine look-up-table outputs to calculate $\mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}'(x'_S, M=2)]$.
Evaluate the Direct-Beam-Emulator to calculate $\mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}'(x'_S, M=1)]$.
Set $\mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}'(x'_S, M=0)] = \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}'(x'_S, M=1)] + \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}'(x'_S, M=2)]$.

percentage error and is, therefore, considered superior. Figure 4.13 shows a comparison between the adjusted expectation of the isotropic and non-isotropic approaches. Once again the result is a large drop in the percentage error. However, both plots show remaining entity dependent error which will be addressed with the residual process $\mathcal{W}^a$.

Building $\mathcal{W}^a$

The final challenge was to update beliefs about the residual process $\mathcal{W}^a$ from equation 4.3.10. To do this a set of pencil beam simulations with a more complex entity were used. A diagram of the entity design is shown in Figure 4.14. r_1, r_2 and the position of the source are varied such that X is well represented. A set of 42 simulations were chosen to train the residual process for this example, shown in Figure 4.15. The set of entities used

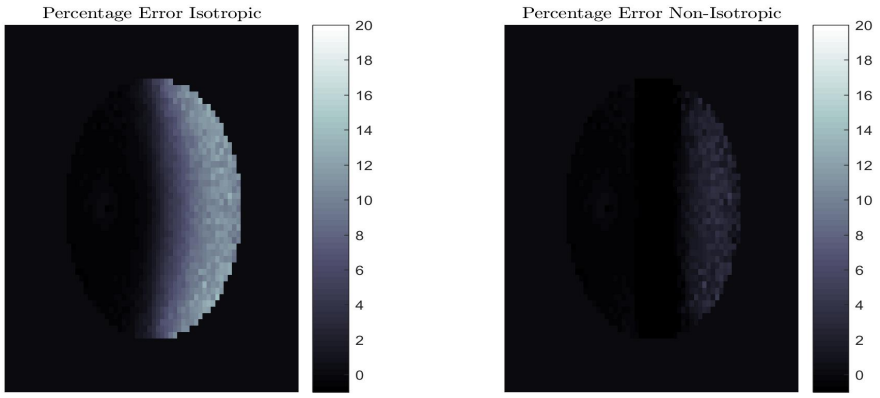

Figure 4.12: Left: percentage error between a pencil beam simulation and the isotropic emulator. Right: percentage error between a pencil beam simulation and the non-isotropic emulator.

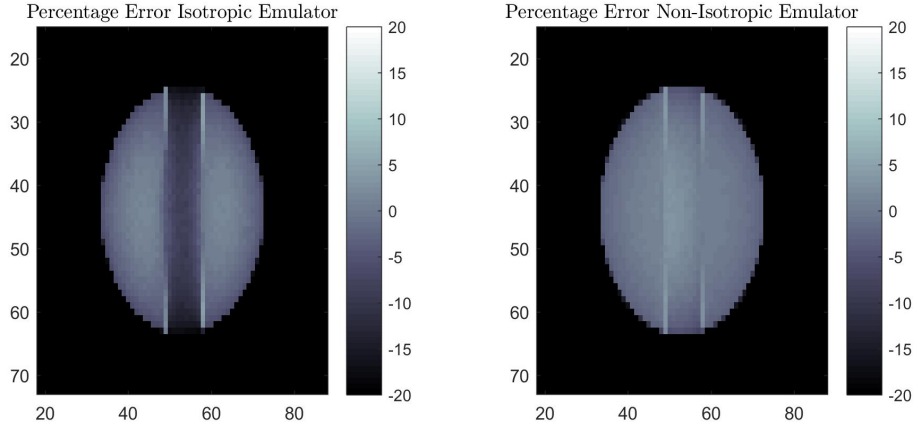

Figure 4.13: Left: percentage error between simulator and isotropic emulator for validation image two. Right: percentage error between simulator and non-isotropic emulator for validation image two.

to train the residual process is notated $\{S^{G4(i)}\}_{i=1:42}$ and the source locations $\{t^{(i)}\}_{i=1:42}$.

For each entity, the quantities $x_S^{(i)}$ and the adjusted expectation of $\mathcal{F}'(x_S^{(i)}, M = t^{(i)})$ are calculated using the look-up-table for every pixel of every training simulation. Then the residual process training-set is

$$X_1 = \{x_a^{(i)}(t^{(i)}), u^{(i,j)}, |t^{(i)} - t^{(j)}|\}_{i=1:42, j=1:\theta\kappa} \text{ and } E = \{e^{(i)}(t^{(l)})\}_{i=1:42, j=1:\theta\kappa}$$

where

$$e^{(i)}(t^{(j)}) = \log\left(\mathcal{F}_{(t^{(l)})}(S^{G4(i)}, M = t^{(i)}) - \exp(\mathcal{LU}(x_a^{(i)}(t^{(j)}), u^{(i,j)}, |t^{(j)} - t^{(i)}|))\right).$$

Here $\mathcal{LU}(\cdot)$ is an evaluation of the look-up-table of adjusted expectations. There is a problem with the log transformation in that we are essentially assuming that $\exp(\mathbb{E}_z[\log(X)]) =$

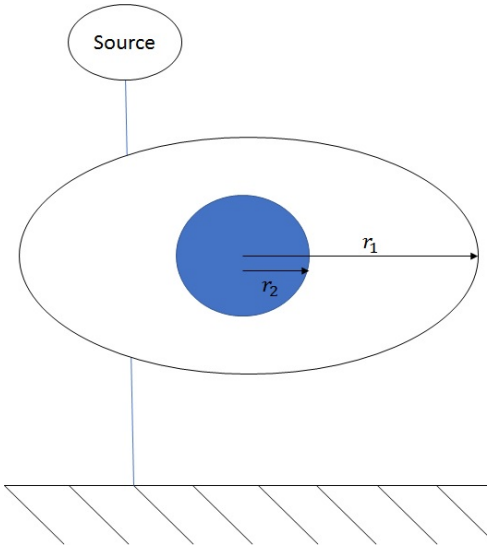

Figure 4.14: Diagram of entity design used to train the residual process. Blue is material one and white is material two. The position of the source and parameters r_1 and r_2 are varied to create the training-set.

$\mathbb{E}_z[X]$. Although not generally true, the error this assumption induces was found to be small enough that it does not have a large impact on the inferences performed. This has been shown on many examples, both in simulation and the real world, which validates the method is robust to the error caused by this assumption.

To make the process cheaper and tractable, a sample of 5000 points was taken from $\{X_1, E\}$ to act as a training-set for the emulator. The method was tested with 1000 points and 10000 points. 1000 resolved much less variation and 10000 did not resolve much more than 5000. Then the residual process is adjusted assuming

$$e^{(i)}(t^{(j)}) = \mathcal{W}^{(a)}(x_a^{(i)}(t^{(j)}), u^{(i,j)}, |t^{(j)} - t^{(i)}|) \oplus \epsilon(x_a^{(i)}(t^{(j)}), u^{(i,j)}, |t^{(j)} - t^{(i)}|)$$

where both $\mathcal{W}^{(a)}$ and ϵ have expectation zero. The ratio of uncertainty was specified directly as

$$\frac{\mathbb{V}ar[\epsilon(x_a^{(i)}(t^{(j)}), u^{(i,j)}, |t^{(j)} - t^{(i)}|)]}{\mathbb{V}ar[\mathcal{W}^{(a)}(x_a^{(i)}(t^{(j)}), u^{(i,j)}, |t^{(j)} - t^{(i)}|)]} = \frac{a_1}{\exp(\mathcal{LU}^{\mathbb{E}}(x_a^{(i)}(t^{(j)}), u^{(i,j)}, |t^{(j)} - t^{(i)}|))} + b_1$$

where $a_1 = 0.001$ and $b_1 = 0.001$ for the example presented in this section. The correlation of $\log(\mathcal{W}^a)$ was specified using equation 2.2.8 with $l_3 = [2, 0.2, 2, 0.2, 32]$ and $p = 2$. The approximated adjusted beliefs were then calculated using the GLS estimates given in key-formula 2.2.1. Figure 4.16 shows the standardised residuals for 500 validation points sampled from E. There were a few points which had high standardised residuals

which indicated an underestimated uncertainty but overall it was considered sufficient. $R^2 = \left(\frac{\mathbb{V}ar[\tilde{E}]}{\mathbb{V}ar[\tilde{E}-\mathbb{E}_E[\tilde{E}]]}\right)^{-1} = 0.83$, where $\tilde{E}$ is a set of 500 validation points, implies that a large portion of the remaining variance was resolved by $\mathcal{W}^a$. To use this technique in practice, a second look-up table is built with the same design as the first. Then the exponential of both look-up-tables are added together and logged so, although there is an increased offline cost, the online cost of the algorithm is still the same. The algorithm for building the residual process look-up-table is given in Algorithm 5.

Algorithm 5 Build $\mathcal{W}^{(a)}$ Look-up-table.

Inputs: $\{S^{G4(i)}\}_{i=1:\mathbb{N}_2}$, $\mathcal{LU}$, a_1, b_1, l_3

Outputs: $\mathcal{LU}$

1: **begin:** Evaluate pencil beam simulations and calculate E and X_1.
2: **for** i in 1:$\mathbb{N}_2$ **do**
3: Evaluate $\mathcal{D}'(\mathcal{D}(S^{G4(i)}))$.
4: Evaluate $\mathcal{F}_{(t^{(l)})}(S^{G4(i)}, M = t^{(i)})$.
5: **for** j in 1:$\theta\kappa$ **do**
6: Evaluate

$$e^{(i)}(t^{(j)}) = \log\left(\mathcal{F}_{(t^{(j)})}(S^{G4(i)}, M = t^{(i)}) - \exp(\mathcal{LU}(x_a^{(i)}(t^{(j)}), u^{(j,i)}, |t^{(j)} - t^{(i)}|))\right)$$

7: **end for**
8: **end for**
9: **end**
10: Make a sample of $\mathbb{N}_3$ points from E and build the residual process emulator using a_1, b_1 and l_3.
11: **begin:** Build residual process look-up table.
12: **for** i in 1:nrow($X^{\mathcal{LU}}$) **do**
13: Evaluate $\mathbb{E}_{\{D^{\mathcal{F}},E\}}[\mathcal{W}^a(x_a^{(i)}(t), u^{(i)}, r^{(i)})]$.
14: **end for**
15: **end**
16: Combine $\mathcal{LU}$ and residual process look-up-table to create new $\mathcal{LU}$.

The summary of the non-stationary-non-isotropic hierarchical-superposition approach to emulating G4 is given by Algorithm 6 and 7.

Algorithm 6 Build Hierarchical-superposition Emulators.

Inputs: $\{S^{G4(i)}\}_{i=1:\mathbb{N}_2}$, a_1, b_1, l_3, X, $X^{\mathcal{LU}}$, a, b, l, l_1, l_2

Outputs: $\mathcal{LU}$ and Direct-Beam-Emulator.

1: Apply Algorithm 3.

2: Apply Algorithm 5.

Algorithm 7 Predict $\mathcal{F}'$.

Inputs: $\mathcal{LU}$, S', Direct-Beam-Emulator

Outputs: $\mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}'(x'_S, M = 0)]$

1: Apply Algorithm 4.

Hierarchical-superposition Results

Figure 4.17 shows the percentage error improvement for the validation points. The overall error has dropped significantly and is now consistently under ±5%. Figure 4.18 shows a comparison between before and after the residual process was added for an entity, as depicted in Figure 4.14, with $r_1 = 1$ and $r_2 = 4$. The error increases for the more complex entity but it is still acceptably low when the adjusted residual process is used.

4.3.5 Uncertainty Estimation

Uncertainty estimation for this technique is technically easy but computationally difficult. The problem is that there are a vast number of adjusted covariances to be considered. For example,

$$\mathbb{V}ar_{D^{\mathcal{F}}}[\mathcal{F}'_{(t)}(x_S, M = 2)] = \sum_{k=1}^{\theta\kappa} \mathbb{V}ar_{D^{\mathcal{F}}}[\mathcal{F}'_{(t)}(x_S, M = t^{(k)})]$$
$$+ 2\sum_{k<l} \mathbb{C}ov_{D^{\mathcal{F}}}[\mathcal{F}'_{(t)}(x_S, M = t^{(k)}), \mathcal{F}'_{(t)}(x_S, M = t^{(k)})]$$

where $\{t^{(i)}\}_{i=1:\kappa\theta}$ is the set of pixels. This means that $0.5\kappa\theta(\kappa\theta - 1)$ adjusted covariances are required for this calculation alone. Therefore, to calculate the adjusted variance at every location requires $0.5(\kappa\theta)^2(\kappa\theta - 1)$ adjusted covariance calculations. The look-up-table approach makes the algorithm feasible but including a look-up-table of adjusted covariances would be computationally impractical as a huge amount of virtual memory would be required to store the table. The result is that the number of points in the adjusted expectation and variance look-up-tables would have to be reduced. This means

that accuracy would suffer in that the error between the approximate adjusted expectation and G4 would increase. As the adjusted covariance look-up-table would require more dimensions, it is also questionable whether a look-up-table would produce a good enough approximation of the adjusted covariances. The choice is whether it is better to have a more accurate emulator, but the assessment of how accurate is poor, or a less accurate emulator with a possibly better assessment of how accurate it is. As key consideration of this study was real-time feasibility, the approach we took was to use the adjusted expectation of the emulator as a *de-facto* deterministic simulator for the rest of the algorithm. As the approximate uncertainty judgements were unsatisfactory and the full uncertainty judgements were too expensive to calculate, this was deemed a reasonable compromise. There is another choice which was briefly explored; a more accurate emulator with an approximation of uncertainty using expert judgement. For example, the adjusted variance is chosen so that there is a 10% error. Although this is not the route taken in this thesis, it did look promising and would constitute an interesting extension to this work.

4.4 Conclusions

In this chapter, a method for building a statistical model of a simulator has been presented. The hierarchical-superposition emulation strategy is an efficient method to emulate simulators with high dimensional inputs and outputs. It combines ideas from signal processing and Bayesian emulation exploiting the advantages of both disciplines. Look-up-tables enable a large number of emulator evaluations to be performed in a reasonable time-frame and inference to be performed.

The application presented in this chapter was emulating G4. The aim was to predict the intensity present in X-ray images caused by photons which have undergone a scattering interaction. The non-stationary-non-isotropic approach outperformed current state-of-the-art approaches by relaxing the isotropic assumptions. The error is small enough that, if G4 is reality and entity composition is known, image quality can be improved over that of an anti-scatter-grid. The online cost of the algorithm is dominated by $\mathcal{D}'(x'_S)$ and $\mathcal{LU}(\mathcal{D}'(x'_S))$. Both of these functions involve a large number operations, essentially $(\kappa\theta)^2$ evaluations of a function. The positive is that the problem is embarrassingly paralisable so large benefits can be gained with GPU computing. Therefore, although it is a computationally intensive method, it can be made fast enough not to impede clinical practice.

There are areas of this method which may be profitable for further research. The first is extending the approach to include a three dimensional model of the entity. This would enable a less heuristic ray-trace procedure, the error induced by layer ordering to be reduced and the method to be extended to computed-tomography applications. There is likely to be a large computational cost with this extension but, if it can be made efficient, it is likely to be superior. The second is the uncertainty estimation. The computational challenges meant that we decided to use the emulator as the *de-facto* deterministic simulator for subsequent analysis. However, other techniques to approximate the adjusted covariance may provide uncertainty estimation which is better than ignoring it.

This chapter has presented a solution to problem statement 1.1.2. However, as G4 is not reality and an entity's composition is unlikely to be known, the benefits of scatter removal cannot be made available to clinicians. The next chapter will consider resolving uncertainty in the entity's composition and Chapter 6 will consider the discrepancy between G4 and reality.

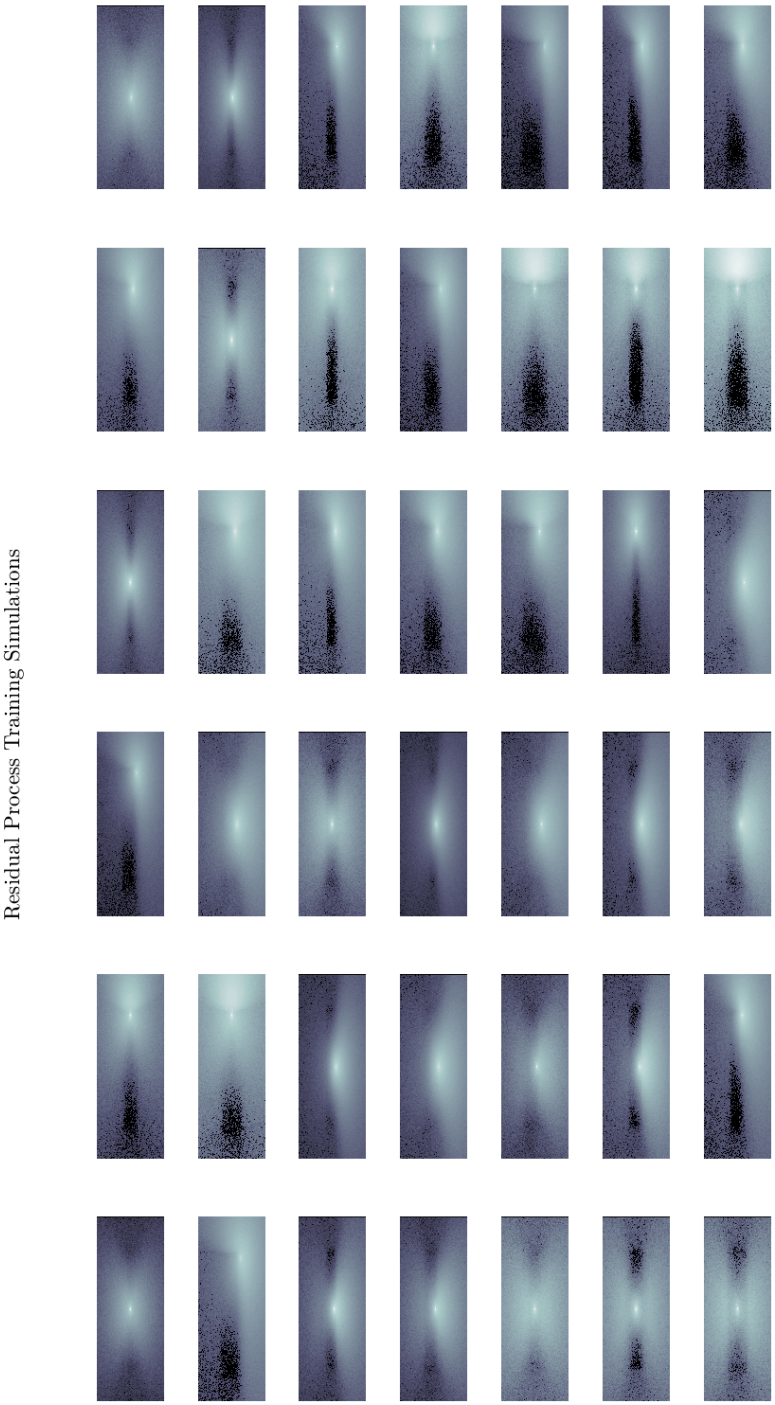

Figure 4.15: Images of all 42 simulations used to train the residual process $\mathcal{W}^a$.

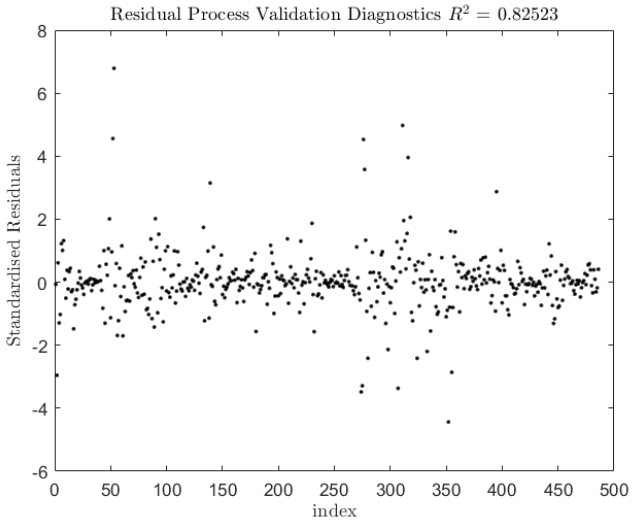

Figure 4.16: Standardised residuals for validation points of the residual process $\mathcal{W}^a$.

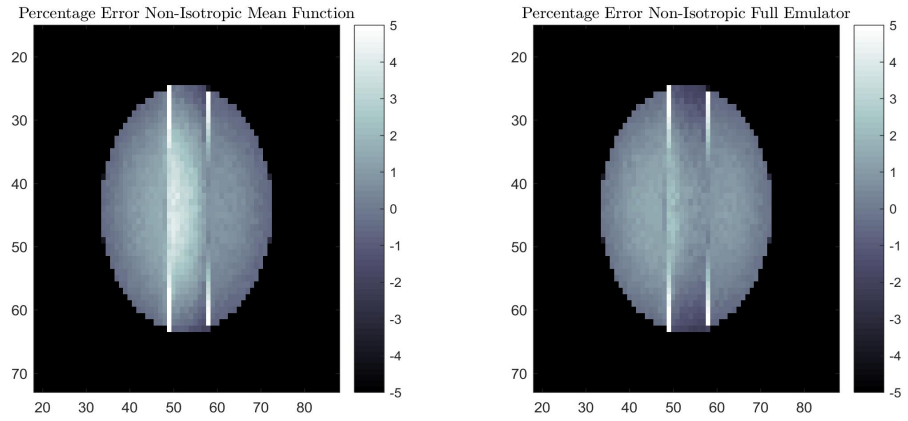

Figure 4.17: Left: percentage error between simulator and non-isotropic emulator without adjusted residual process. Right: percentage error between simulator and emulator with adjusted residual process.

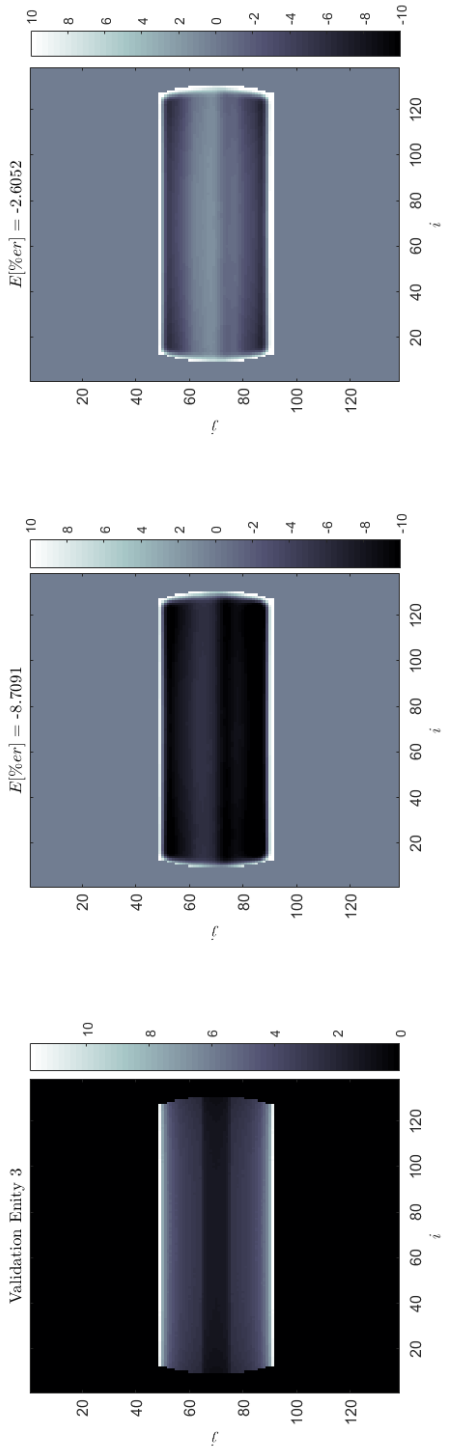

Figure 4.18: Left: simulator output for validation entity 3. Centre: percentage error between simulator and non-isotropic emulator without residual process. Right: percentage error between simulator and non-isotropic emulator with residual process.

Chapter 5

History Matching and Random Field Priors

In the previous chapter, how to predict an X-ray image given an entity's composition was presented. However, an entity's composition is often unknown. If the composition of a patient can be inferred from an ordinary X-ray image, the effect of scattered photons may be reduced, improving image quality. Making inferences about the composition of an entity can also be useful to clinicians for diagnosing particular conditions. For example, a measure of areal Bone Mineral Density (aBMD) can inform whether a patient has osteoporosis. Therefore, the task of composition inference has two purposes: i) infer diagnostic quantities and ii) reduce uncertainty in the scatter present in an X-ray image and, therefore, a necessary component of a scatter removal method (SRM).

This chapter will present the use of i) the Bayes linear approach to making inferences about multivariate random fields (presented in Chapter 2) and ii) the emulator of an X-ray examination system (presented in Chapter 4) to making inferences about an entity's composition from an X-ray image of the entity. Here composition inference is formulated as an inverse problem, meaning, given an observation of system behaviour, adjust an expert's beliefs about the system values. Here the i) system is a piece of X-ray examination equipment, ii) system behaviour is an X-ray image and iii) system values are the composition of the entity being examined. Beliefs about both system values and behaviour may therefore be organised using multivariate random fields. The solution must be efficient as clinical practice can not be impeded by the amount of time it takes to calculate adjusted beliefs. Therefore, in order to solve problem statement 1.1.3, an efficient methodology is required for solving inverse problems where system behaviour and values are described by

multivariate random fields.

Inverse problem solving is not a new concept in X-ray imaging [79]. It is almost exclusively applied to computed-tomography (CT) where the aim is to reconstruct a three dimensional representation of the entity being examined from multiple X-ray images of the entity. Methods include i) back propagation [79], ii) regularised maximum likelihood [80], iii) neural networks [81] and iv) Bayesian approaches [82, 83]. The work in this thesis differs from these approaches to inverse problem solving with X-ray images in three ways. The first is that all approaches reviewed use multiple images to try and make inferences about the entity being examined. The second is that we use GEANT4 as the forward model and their forward model makes a linear approximation of the system. This is essentially assuming that the photons are mono-chromatic (see equation 3.5.1) and there is no scatter. These assumptions cause two problems: i) as the beams are poly-chromatic (see equation 3.5.2) and there is scatter, an error is induced and ii) if the assumptions are correct, little uncertainty in the entity's composition can be resolved using a single X-ray image. We believe the ubiquity of this linear simplification has prevented inverse problem solving approaches using a single X-ray image. If scatter is considered, as will be demonstrated in this chapter, useful amounts of uncertainty in an entity's composition can be resolved. The third difference is that the inverse problem is solved using the history matching approach. This approach reduces some of the practical problems with Bayesian inverse problem solving, namely, i) prior elicitation, ii) computation of the posterior and iii) robustness [32,35]. Therefore, our methods differ from current approaches as, i) inferences are made using a single X-ray image, ii) our forward model is more realistic and iii) our Bayesian inverse problem solving approaches can return inferences in a useful time-frame.

This chapter will be organised as follows: Section 5.1 formulates inverse problems and current approaches, Section 5.2 describes the history matching approach to inverse problem solving, Section 5.3 extends the history matching approach to include random field priors, Section 5.4 proposes an iterative approach to finding non-implausible points and Section 5.6.1 provides a proof-of-concept in simulation of our solution to problem statement 1.1.3.

5.1 Inverse Problem Solving

The relationship between system values and system behaviour is given by

$$y = \mathcal{F}(x^\star) + \epsilon_{md} \text{ and} \tag{5.1.1}$$

$$z = y \oplus \epsilon_{me} \tag{5.1.2}$$

where $\mathcal{F}$ is the simulator mapping system values to system behaviour, $x^\star$ is a vector of the system values considered the causal factors of system behaviour, y is system behaviour, ϵ_{md} is the model discrepancy, z is an observation of system behaviour and ϵ_{me} describes the measurement error when the observation was made. Making inferences about $x^\star$ is known as the inverse problem. As the aim of this chapter is to present a proof-of-concept of problem statement 1.1.3 in simulation, it will be assumed that $\epsilon_{md} = 0$. A proof-of-concept in simulation is an important proof point on the path to a software solution to the scattering problem. It does not show that the method will transfer to the real world, however, if a method does not work when $\epsilon_{md} = 0$, it is unlikely to work when $\epsilon_{md} \neq 0$. Inverse problem solving has been applied to a number of physical systems: for example, oil reservoir management [27,34,35], galaxy formation [32], climate change [22,41], disease modelling [36] and systems biology [37].

In this chapter, there are two goals of inverse problem solving which are of interest. The first is when the system values are of interest so resolving uncertainty in their value directly enables the goals of a study to be met see, for example, [32, 35, 37, 41]. The second is when the system values are of interest as knowledge of their value would resolve uncertainty in system behaviour see, for example, [22, 27, 34]. In this thesis, both are of interest. Reducing uncertainty in the system values means that i) diagnostic measures can be given to clinicians and ii) uncertainty about the intensity due to scatter is reduced. Therefore, by resolving uncertainty in an entity's composition, we can predict scatter in an X-ray image more accurately, improving image quality, and return measures which can be used by clinicians to diagnose diseases.

Inverse problems are difficult to solve when they contain at least one of the following features.

1. They may be under-determined in that the system values are higher dimensional than the observations. This implies the system values may not be identified precisely.

2. $\mathcal{F}$ may not define an injection. This means that there are multiple values of x which

match system behaviour equally well, but, they are not all related to system values.

3. System behaviour may be uncertain as it cannot be observed precisely. Then, even if there is a unique inverse of $\mathcal{F}$, the system values may not be identified precisely.

4. The forward operator is complex and expensive to evaluate, therefore, it is only known at the locations at which the simulator has been evaluated.

5. The input dimension may be large making it challenging to effectively explore the space of possible system values.

Given these constraints, an approach based solely on optimisation − that is maximum likelihood or minimising an error to observed system behaviour − to this problem may be impracticable as i) a unique solution may not exist, ii) if it does exist, due to measurement error it may not be the system values and iii), if there is no measurement error, due to the size of the space and the cost of the forward operator, it may not be possible to find it. For example, if the aim is to make inference for (μ, η) using a single pixel X-ray examination system and the Beer-Lambert law (see equation 3.5.2), i) (μ, η) are in-separable so a unique solution does not exist, ii) the measurement will contain Poisson noise and iii), if any thickness and alloy is possible, the solution space is high dimensional. A solution to these problems is the Bayesian approach to inverse problem solving. Optimisation procedures may still be useful, for example by identifying an interesting region of the input space to perform a Bayesian analysis, but they do not always provide a solution to the inverse problem which can be used to make the required decisions.

5.1.1 Bayesian Inverse Problem Solving

In the full Bayes solution to inverse problems, $\{y, z, x^\star\}$ are random quantities. Then the solution is defined by the posterior distribution $\mathbb{P}(x^\star|z)$ which is found using Bayes' theorem $\mathbb{P}(x^\star|z) \propto \mathbb{P}(x^\star)\mathbb{P}(z|x^\star)$.

In the Bayesian approach, the belief that some solutions are improbable can be coherently included in the inference. Then no matter how small the error between the simulator output and an observation, these solutions will not be considered probable or acceptable. To apply the full Bayes approach, the specifications required are a likelihood and a prior distribution for $x^\star$. As the aim is to calculate an expert's posterior distribution, the Bayesian solution is not ill-posed. This is not to say that the posterior is useful. For

example in X-ray imaging, if the aim is to infer the composition of an entity using only direct beam, the posterior credible set may not resolve any uncertainty about the material present. The full Bayes solution to inverse problem solving is difficult to implement for a number of reasons:

1. as a prior distribution for $x^\star$ must be specified, $x^\star$ must have a clear physical meaning,
2. $x^\star$ may be high dimensional making meaningful full distributional specifications challenging to elicit,
3. Monte Carlo sampling algorithms are often required to compute the posterior and
4. the problem of an expensive simulator is not overcome.

An alternative to the full Bayes approach is Bayes linear methods, see Section 2.1.1. The prior specification is more reasonable and the calculations are cheaper. However, in order to calculate $\mathbb{E}[z], \mathbb{V}ar[z]$ and $\mathbb{C}ov[x^\star, \mathcal{F}(x^\star)]$, a large number of evaluations of the simulator are required and a single covariance matrix may not be sufficient to describe the complex relationship between system values and behaviour. Furthermore, if $x^\star$ and y are high dimensional, the covariance matrix $\mathbb{C}ov[x^\star, z]$ can be too large to store in memory.

5.1.2 The Emulation Approach

The emulation approach incorporates a statistical model to organise beliefs about the forward operator $\mathcal{F}$. Then y, z, $x^\star$ and $\mathcal{F}$ are all considered uncertain. As discussed in the previous chapter, emulators are usually cheaper to evaluate and, as such, may be used to reduce the computational complexity of inverse problem solving. An emulator allows the uncertainty in $\mathcal{F}$ to be integrated out and inference is performed using $\mathbb{P}(x^\star|z, D^{\mathcal{F}})$ [22]. Emulation has made great strides in tackling a large number of previously intractable problems see, for example, [31, 38]. However, the efficiency of the inference is still limited by the cost of Monte Carlo sampling algorithms. If the output is not available in a useful time-frame using full Bayes, a more efficient method is required.

5.2 History Matching for Inverse Problem Solving

History matching is an alternative subjective Bayesian method for inverse problem solving. The term originated in the oil industry, an early application of the method [35]. Now it often refers to a tool for Bayesian inverse problem solving where the aim is to find a set

that is considered as the set of inputs which return an acceptable match to observed data given the uncertainties in the system. The process simply finds acceptable matches to observed data. It is not required that an acceptable match is related to a physical entity which may be measured independently of the simulator. The set of acceptable matches may be associated with a physical quantity, but, that is not built into the base assumptions of the method. Therefore, if history matching is to be used to make inferences about a system value, another layer of modelling is required and justification that the acceptable matches are in fact related to a physical quantity. This is not a limitation of the approach; it gives the analyst more freedom using the subjective Bayesian paradigm, as analysis may be performed on quantities which have an ambiguous operational definition. For example, a simulator's input may be too under-resolved to describe system values meaningfully. Then it is questionable whether there is a single $x^\star$ which describes system values. History matching does not assert that a single $x^\star$ exists and only returns the set of xs such that if $x^\star = x$, an acceptable match is attained.

History matching is performed using implausibility measures where large implausibility usually indicates that x does not give an acceptable match to observed data. The practical implementation of history matching is attractively simple.

1. The process begins by setting prior bounds for the value of $x^\star$ and any point within the bounds is referred to as non-implausible.

2. The space − referred to as the non-implausible space − defined by these bounds is sampled using, for example, a latin-hyper-cube [32]. Then for every point in the sample, the simulator is evaluated − or beliefs about the simulator adjusted − and an implausibility measure is calculated.

3. A threshold is set and, for any points with implausibility greater than the threshold, it is stated that there is evidence that the simulator evaluated at this input is unlikely to provide an acceptable match to observed data. These points are referred to as implausible and the set which is not declared implausible is referred to as the posterior non-implausible set.

History matching aims to find a set of acceptable matches to observed data, referred to as non-implausible points. The best way to ascertain if any non-implausible points exist is to test a large enough number of points so that every point has been effectively tested. This is referred to as an exhaustive search. Given that an exhaustive search is expensive

when the simulator is expensive to evaluate, an emulator is used instead. The emulator enables implausibility measures to be computed for a huge number of points and therefore makes an exhaustive search feasible.

Implausibility Measures

Implausibility measures are based on comparing the observation and the output of a simulator scaled by uncertainty. They are an artefact which organises beliefs about whether an acceptable match to observed data has been attained. There are three implausibility measures which are used in this thesis, uni-variate implausibility, max implausibility and mean implausibility.

1. The first measure is a uni-variate implausibility measure

$$\mathcal{IM}(z(t),x) = \left| \frac{z(t) - \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}_{(t)}(x)]}{\sqrt{\mathbb{V}ar_{D^{\mathcal{F}}}[\mathcal{F}_{(t)}(x)] + \mathbb{V}ar[\epsilon_{me}(t)]}} \right|. \tag{5.2.1}$$

   To decide if a point is implausible, Pukelsheim's three sigma rule is used [47]. For example, if $x^\star = x$, the probability of an implausibility greater than three is less than 0.05 so it is unlikely that $x^\star = x$ will provide an acceptable match.

2. The second measure is the multivariate extension to the first

$$\mathcal{IM}^2(z,x) = (z - \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}(x)])^T (\mathbb{V}ar_{D^{\mathcal{F}}}[\mathcal{F}(x)] + \mathbb{V}ar[\epsilon_{me}])^{-1} (z - \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}(x)]). \tag{5.2.2}$$

   The expected value of $\mathcal{IM}^2(z,x) = \text{rank}(\mathbb{V}ar[\mathcal{F}(x)])$ and the variance depends on the distribution of z. For example, if Gaussian assumptions are made for simulator and measurement error, $\mathcal{IM}^2(z,x)$ is chi-squared distributed. The advantage of this measure is that it includes the correlation structure. The disadvantages are that the decision is complicated by dependence on possibly unspecified distributions and for large outputs the inversion can be challenging.

3. The third measure takes the maximum over the uni-variate measures

$$\mathcal{IM}^{\max}(z,x) = \max_t \left\{ \left| \frac{z(t) - \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}_{(t)}(x)]}{\sqrt{\mathbb{V}ar_{D^{\mathcal{F}}}[\mathcal{F}_{(t)}(x)] + \mathbb{V}ar[\epsilon_{me}(t)]}} \right| \right\}. \tag{5.2.3}$$

   This is a conservative multivariate measure. It is specifying that if an expert believes that any output does not match observed behaviour, then the input is implausible. There are two criticisms of this measure. The first is that it does not account for

the correlation structure. The second is that if it is assumed that the outputs are independent and the observation vector is large, it is likely that it will be classified as implausible by chance.

4. The final measure is the mean of the uni-variate implausibility measures. This is stating that if on average it is acceptable, then it is acceptable. It is the weakest criteria in that it favours the non-implausible decision least but it is reasonable for large observation vectors. There are two prominent criticisms of this measure. The first is the correlation structure is not taken into account. The second is that every output having uncertainty close to the threshold can be considered better than all outputs having low implausibility and one output having a very large implausibility. This can be checked by the analyst using diagnostic plots or, as in the example presented at the end of this chapter, diagnostic images. Visual diagnostics are not ideal when providing software for use by non-statistical experts but better than no diagnostics at all.

Toy Example: History Matching

To illustrate the history matching approach, the toy example from Section 4.2 is used again. Figure 5.1 shows the history matching procedure applied to the toy example of a deterministic simulator. The uni-variate implausibility measure was used and the implausibility threshold is three. Where the uncertainty bounds of simulator and observation overlap, the implausibility is below the threshold as there is no evidence that those points did not produce the observation.

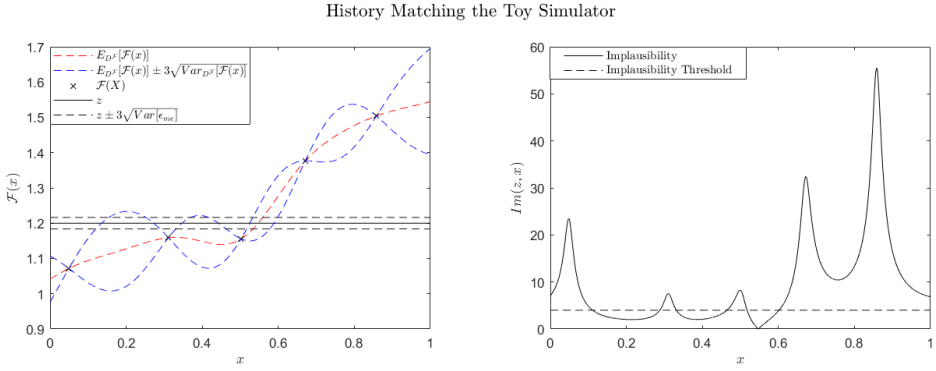

Figure 5.1: Left: adjusted beliefs about the toy simulator and a toy observation. Right: implausibility as a function of x.

History Matching in Waves

When the uncertainty in the simulator is not sufficiently resolved by the emulator, history matching may be performed in waves. A wave starts with a non-implausible set. From this set, an emulator training-set is designed and the simulator is evaluated at the points in the design. An emulator is built and the history match performed, creating a refocused non-implausible set. This process continues until the non-implausible set stops changing between waves. The wave approach has two advantages: i) it is computationally more efficient for large spaces and ii) the emulator is better able to describe the simulator in a local region in later waves. The analysis for wave i, with non-implausible set $\chi^{(i)}$, is as follows:

1. choose a set of points in $\chi^{(i)}$,
2. evaluate the simulator at those points creating data-set $D^{\mathcal{F}(i)}$,
3. adjust beliefs about the simulator using $D^{\mathcal{F}(i)}$,
4. calculate implausibility for a large sample of points in $\chi^{(i)}$ and
5. remove all points above a threshold to produce $\chi^{(i+1)}$.
6. If $\chi^{(i+1)}$ is approximately the same size as $\chi^{(i)}$ stop,
7. otherwise return to step 1.

When the input dimension is large and the simulator expensive, wave based history matching has made significant inroads into previously intractable problems. It is computationally efficient and overcomes many problems prominent in Bayesian calibration. For example, a valid emulator may not be built for all outputs initially but it can when the space of acceptable matches is narrowed down. Then the outputs which can be emulated initially can be used to narrow the space and additional outputs can be added in later waves. This is referred to as sequential history matching [37]. A wave based approach struggles when the input dimension is large. This is because making enough samples to densely represent the prior non-implausible space is difficult. For example, suppose a Latin-hyper cube is used to sample the prior non-implausible space. The posterior non-implausible space can be small enough that, even if the simulator is free to evaluate, sampling enough points to well represent the posterior non-implausible set can be impractical. Other methods, like splice sampling or a simple MCMC have been proposed,

see [36, 37], but they have not yet made it into common usage. Other sampling regimes may be possible but research into these are in their infancy. Another problem with the emulation approach, caused by high dimensionality and a rough simulator, is that the residual process has essentially no effect as the input space is only sparsely represented. Therefore, the emulator is essentially a linear model. If the output is rough, the linear model does not resolve much uncertainty and the amount of space cut out in each wave is small.

The emulator used in the examples in this thesis is not cheap to evaluate by usual standards. This is because it requires $(\theta\kappa)^2$ evaluations of a ray-trace algorithm and $(\theta\kappa)^2 + \theta\kappa$ evaluations of an emulator. There were a couple of natural potential avenues to dealing with our problem which were not taken. The first is to sequentially history match. This is when outputs are sequentially used to reduce the non-implausible space. The problem with this approach is that each output requires $(\theta\kappa)$ ray traces and $(\theta\kappa) + 1$ emulator evaluations. This is still an expensive piece of computer code and the problems with sampling the $(2\theta\kappa)$ dimensional space remain. The dimension $(2\theta\kappa)$ is misleading as the correlation in the inputs imply a much smaller effective dimension, see Section 5.3. For example, a spatial model of a NOF could be designed with far fewer parameters than $(2\theta\kappa)$. Therefore, a combination of i) a spatial model of a NOF, ii) an emulator of the emulator in the previous section and iii) sequential history matching, is a natural route to go down. This route was not chosen because the number of parameters required to describe a NOF would be large. This means the problems using an online wave based history match still remain, namely i) the effective simulator is too expensive to evaluate online a sufficient number of times, ii) sampling the space is still troublesome and iii) validating and evaluating many emulators is challenging to produce in software intended for use by non-statistical experts. Therefore, as the emulator of G4 is expensive and the size of the input space is large, wave based history matching does not play a prominent role in the rest of this thesis. It is, however, an excellent tool to use for the type of problems presented. See [35–37,39] for further details on the wave based history matching approach.

5.2.1 History Matching in Subjective Bayesian Analysis

History matching is an effective method for solving inverse problems with complex computer models. It is now considered whether it has a more general role in Bayesian analysis. The first situation considered is using history matching to augment the Bayes linear ap-

proach. A situation which is problematic when using Bayes linear is when a problem has two uncertain quantities and adjusting y given $x^\star$ is simple but adjusting $x^\star$ given y (or both together) is hard. Difficulties in adjusting beliefs about a random variable occur when i) the covariance between a quantity of interest and the data is *small* or ii) the relationship is nonlinear. In both these scenarios, the Bayes linear adjustment is a sub-optimal paradigm as the second order specification is not flexible enough to encode an expert's beliefs. Instead of calculating adjusted expectations and variances directly, an additional stage can be added to find the non-implausible set and an expert's beliefs are then adjusted using that set. For example, consider an observation modelled by a mixture of random fields where the partitions are defined by uncertain control points, see Section 2.3. An additional stage can be added to the inference which removes control point combinations which are highly unlikely to have produced the observed data. Then, if the control point themselves are of interest, beliefs can be adjusted by the non-implausible sets or, if prediction at a location on a field is of interest, beliefs are adjusted by averaging over the non-implausible set.

This situation is addressed using a method constructed to be analogous to inverse problem solving with complex computer models. The process assumes that adjusting beliefs about $x^\star$ is challenging – as the specifications are hard and linear fitting does not resolve sufficient uncertainty – but adjusting beliefs about y given $x^\star = x$ is tractable. The relationship between observations and unknown quantities is given by

$$\begin{aligned} z &= y \oplus \epsilon_{me} \text{ where} \\ y &= \mathcal{F}(x^\star) \oplus \mathbb{A}_z[y] \text{ and} \\ \mathcal{F}(x) &= \mathbb{E}_z[y; x^\star = x]. \end{aligned} \tag{5.2.4}$$

Here $\mathbb{A}_z[y]$ is the adjusted version of y by z, see definition 2.1.7. If $x^\star = x$ is inconsistent with the uncertainties expressed in equation 5.2.4, it is judged that it should not play any further part in the inference. The rationale is that given $x^\star = x$, $\mathbb{P}(\mathcal{IM}(z, x) > 3) < 0.05$ so it is unlikely that $x^\star = x$ or that when $x^\star = x$ a useful model will be forthcoming. Therefore, history matching can be used to find the non-implausible set for $x^\star$, whether x is the input to a simulator, a measurable uncertain quantity or a hyper-parameter in a model.

The analyst is free to decide which features of the modelling assumptions they would like to assess and to design an appropriate implausibility measure to test it. One generic approach is to split the observation vector z into two $\{z(T'), z(T^*)\}$. Beliefs about $z(T^*)$

are adjusted using $z(T')$ and implausibility calculated using, for example

$$\mathcal{IM}(z,x) = \max_{k,t^* \in T^*} \left\{ \left| \frac{z_{(k)}(t^*) - \mathbb{E}_{z_{(k)}(T')}[z(t^*); x^\star = x]}{\sqrt{\mathbb{V}ar[\epsilon_{me(k)}(t^*)] + \mathbb{V}ar_{z(T')}[y_{(k)}(t^*); x^\star = x]}} \right| \right\}$$

where $t^* \in T^*$. This approach circumvents the difficulties of adjusting $x^\star$ directly by adjusting y given $x^\star = x$. A common method used to validate emulators is analogous to this approach. Adjusting beliefs about the correlation length is challenging but adjusting beliefs about the emulator given a correlation length is tractable. Therefore, it is fixed to a value which passes a set of validation diagnostics, for example, cross validation. This process is analogous to finding a single non-implausible point and using it as a point estimate going forward.

Belief Adjustment with Non-Implausible Sets

If a Bayes linear analysis is being performed, an alternative is to use non-implausible sets in the case where the relationship between $x^\star$ and $\mathcal{F}(x^\star)$ is not well described by a covariance matrix. In this section, a novel method for adjusting beliefs about random quantities with non-implausible sets will be presented. Given the posterior non-implausible set has been found, the next step is to adjust beliefs about $x^\star$ with this set. Here the method is to take a summary statistic from the non-implausible sets and adjust beliefs using that summary statistic. The posterior non-implausible set is notated $\chi = \{x^{(i)}\}_{i=1:\mathbb{D}(\chi)}$ where $\mathbb{D}(\chi)$ is the number of points in χ. Three summary statistics will be discussed: minimum implausibility, mean implausibility and weighted implausibility.

1. Minimum implausibility is when the point estimate is given by

$$\bar{x}^\star = x^{(i)} \text{ such that } \mathcal{IM}(z, x^{(i)}) = \min_j \left\{ \mathcal{IM}(z, x^{(j)}) \right\}$$

   where $x^{(i)}$ notates the vector which is the i^{th} member of the posterior non-implausible set. As low implausibility does not, in general, imply it is more plausible, there is nothing special about the minimum implausible point. However, if a single point is chosen from the set to represent it, unless the expert has specified beliefs about the distributions involved, there is no point that is expected to be better than the minimum implausible one. The advantage of this point estimate is that it is always non-implausible.

2. Mean implausibility is when the point estimate is given by

$$\bar{x}^\star = \frac{1}{\mathbb{D}(\chi)} \sum_{i=1}^{\mathbb{D}(\chi)} x^{(i)}. \tag{5.2.5}$$

The mean is effective but, if the posterior non-implausible sets are non-convex, there is no guarantee that the value itself is non-implausible. It also has problems when the system value is on the boundary of the prior non-implausible set.

3. Weighted implausibility is when the point estimate is given by

$$\bar{x}^\star = \sum_{i=1}^{\mathbb{D}(\chi)} \alpha_{(i)} x^{(i)}$$

where

$$\alpha_{(i)} = \frac{(\mathcal{IM}(z, x^{(i)}))^{-1}}{\sum_{j=1}^{\mathbb{D}(\chi)} (\mathcal{IM}(z, x^{(j)}))^{-1}}.$$

It is a good hybrid which is more robust than mean implausibility as a few points with relatively high implausibility in a second mode will not have a high impact on the inference. It is essentially assuming that low implausibility implies higher plausibility which is not always true.

Once the summary statistic has been chosen, second order specifications for $\{x^\star, \bar{x}^\star\}$ may be made directly or using an automated sampling procedure. As with other automated sampling procedures, these require distributional specifications to be made. As an expert may be unwilling to assert that these distribution are commensurate with their beliefs, a robustness analysis should be performed to ensure that the inferences performed are not unduly sensitive to the assumptions. See Appendix A.3 for an application of robustness analysis to the specifications required for a fourth order analysis. This procedure may be performed before observations have been made as part of the belief elicitation and prior diagnostic procedures. The sampling algorithm is as follows.

1. A sample of points, notated $\mathcal{X}$, is taken from a distribution with expectation and variance matching $\mathbb{E}[x^\star]$ and $\mathbb{V}ar[x^\star]$.

2. $z^{(i)} = \mathcal{F}(x^{(i)}) \oplus \epsilon_{me}^{(i)}$ is sampled for all $x^{(i)} \in \mathcal{X}$.

3. For every point $z^{(i)}$, a non-implausible set is inferred by history matching and the summary statistic calculated.

4. The second order specification for $\{x^\star, \bar{x}^\star\}$ is calculated using the usual unbiased estimators for population mean, variance and covariance.

An important note is that this method is effective when the non-implausible set is uni-modal. If it is not, further work is required to perform a Bayes linear analysis. If $\mathcal{F}$ is expensive to evaluate, an emulator may be used instead and if $x^\star$ is high dimensional, analysis may be performed in waves.

Prediction with Non-Implausible Sets

For prediction, one option is to average over the non-implausible set. The adjusted beliefs are approximated by

$$\mathbb{E}_z[y] \approx \frac{1}{\mathbb{D}(\chi)} \sum_{i=1}^{\mathbb{D}(\chi)} \mathbb{E}_z[y; x^{(i)}] \text{ and}$$
$$\mathbb{V}ar_z[y] \approx \frac{1}{\mathbb{D}(\chi)} \sum_{i=1}^{\mathbb{D}(\chi)} \mathbb{V}ar_z[y; x^{(i)}] + \frac{1}{\mathbb{D}(\chi) - 1} \sum_{i=1}^{\mathbb{D}(\chi)} (\mathbb{E}_z[y; x^{(i)}] - \mathbb{E}_z[y])^2.$$

Another option is to weight the average by implausibility. The adjusted beliefs are approximated by

$$\mathbb{E}_z[y] \approx \sum_{i=1}^{\mathbb{D}(\chi)} \alpha_{(i)} \mathbb{E}_z[y; x^{(i)}] \text{ and}$$
$$\mathbb{V}ar_z[y] \approx \sum_{i=1}^{\mathbb{D}(\chi)} \alpha_{(i)} \mathbb{V}ar_z[y; x^{(i)}] + \sum_{i=1}^{\mathbb{D}(\chi)} \alpha_{(i)} (\mathbb{E}_z[y; x^{(i)}] - \mathbb{E}_z[y])^2$$

where

$$\alpha_{(i)} = \frac{\mathcal{IM}^{-1}(z, x^{(i)})}{\sum_j \mathcal{IM}^{-1}(z, x^{(j)})}.$$

Weighted averaging is appropriate when the prior is considered uniform and the likelihood is uni-modal and symmetric. If this is not the case, small implausibility does not imply that a parameter combination is more probable so weighting in this way is incoherent.

The third option is to use a point estimate, for example the expert's adjusted expectation for the system values or a point from the non-implausible set. The advantage of point estimates is that they are cheaper than approximate integrals. The disadvantage is that additional work is required to propagate the uncertainty in the system value through to uncertainty in system behaviour, see [22] for an approach to performing an analysis of this type.

For the goals of this study, prediction with a point estimate was acceptable. The aim of prediction was to improve image quality by correcting scatter in images and a point estimate was sufficient for this purpose. It was also unclear how accurate uncertainty judgements could be incorporated in the visual output in a way that would be useful to clinicians. This will be discussed in further detail in Chapter 6.

History Matching for Uncertain Look-ups

A common problem in STEM (Science, technology, engineering and maths) research and development is given an observation z, calculate the value of interest $x^\star$. In other words, the aim is to look-up the value associated with an observation from a sparse database. The database is created by making observations at a number of known $x^{(i)}$s, notated $\{X, D^f\} = \{x^{(i)}, z^{(i)}\}_{i=1:N}$ and the expert would like a function which takes an input z and returns the quantity of interest $x^\star$. The simplest approach would be to measure $z^{(i)}$ at enough $x^{(i)}$s that the prior space of $x^\star$ is densely represented and, for a new observation z, return the $x^{(i)}$ which minimises $|z^{(i)} - z|$. This is impractical as making enough observations to densely represent the space is very expensive. It is also incoherent as the observations are imprecisely measured so the point which minimises $|z^{(i)} - z|$ need not be $x^\star$. The next point of call is usually to fit a functional relationship

$$\mathcal{B}(z^{(i)}) \oplus \epsilon^{(i)} = x^{(i)}$$

using a sparse training-set. It is popular because it returns a function with inputs z and outputs $x^\star$ and the tractable least squares methodology may be applied using a few lines of code or using the various toolboxes available. The least squares estimators are calculated and prediction may then be performed for new observations with associated uncertainty judgements. Although it can sometimes perform well, it violates the usual regression assumptions, namely that the independent variable is error free while there is an error between the dependent variable and the ground truth. This approach can therefore perform badly for at least two reasons: i) z is measured imprecisely and ii) the covariance structure going from observations to quantities of interest directly can be too complex. Another approach is to fit z as a function of x and invert the relationship either analytically or numerically to find x. This method struggles with uncertainty judgements and can exhibit the problems associated with minimisation approaches to inverse problems (see Section 5.1). The alternative which is presented in this thesis is to use the history

matching approach. The relationship is given by

$$\begin{aligned} \mathcal{F}(x^{(i)}) &= y^{(i)} \text{ and} \\ z^{(i)} &= y^{(i)} \oplus \epsilon_{me}^{(i)}. \end{aligned} \tag{5.2.6}$$

Beliefs about $\mathcal{F}$ are adjusted using D^f. History matching can be used for the inversion step to get a non-implausible set. This set can be used directly for decision making or beliefs may be adjusted using the non-implausible set. This approach will be used in the industrial example to make inferences about an entity's composition from an X-ray image taken on a Multi Adsorption Plate (MAP) equipped detector.

Belief Adjustment for Multivariate Random Fields by Non-Implausible Sets

Suppose instead of an observation vector, the elements of which are locations on a multivariate random field, a set of non-implausible sets are available where a member of each set corresponds to a location on a multivariate random field. In this section, a novel approach to adjusting beliefs about a multivariate random field with sets of non-implausible sets will be presented. Each non-implausible set is notated $\chi(t) = \{x_{(j)}^{(i)}(t)\}_{i=1:\mathbb{D}(\chi(t)),j=1:p}$. Given that beliefs about system values are specified as a multivariate random field, a method is required to use these beliefs to refocus the non-implausible sets at every location. The approach proposed here refocuses the sets one field at a time.

1. Calculate a summary statistic at every location on the i^{th} constituent field, notated $\bar{x}_{(i)}^{\star}(T) = \{\bar{x}_{(i)}^{\star}(t^{(k)})\}_{k=1:n}$.

2. Assume that

$$\bar{x}_{(i)}^{\star}(t) = x_{(i)}^{\star}(t) \oplus \epsilon(t) \tag{5.2.7}$$

   where $\mathbb{E}[\epsilon(t)] = 0$ and

$$\mathbb{V}ar[\epsilon(t)] = \sigma_{x(i)}(t) = \frac{1}{\mathbb{D}(\chi(t)) - 1} \sum_{i=1}^{\mathbb{D}(\chi(t))} \left(x_{(i)}^{(k)}(t) - \frac{1}{\mathbb{D}(\chi(t))} \sum_{l=1}^{\mathbb{D}(\chi(t))} x_{(i)}^{(l)}(t) \right)^2.$$

   This implies $\bar{x}_{(i)}^{\star}(T)$ is a random vector, the elements of which are locations on a multivariate random field, correlated with the quantity of interest, and beliefs can be adjusted using Section 2.2.2. A major motivator for trying this construction was that it would be in a form that we were comfortable with dealing with, namely a multivariate random field. However, there are other reasons it is a good representation. If a random quantity can take a set of possible values, a second order representation

concise and it is often enough to be able to perform a useful uncertainty analysis. For example, Bayes linear analysis uses this idea and Gaussian approximations are generally popular in statistics. Another motivation is provided by the function space view of Gaussian processes. A sample from a Gaussian process is a function and the posterior can be considered as all sampled functions which intersect the training points [30]. One way of adjusting beliefs using fields of non-implausible sets would be to sample a Gaussian process and return all functions which intersect a non-implausible point at every location. This would be prohibitively expensive but the method presented can be viewed as a fast approximation to this analysis. This analysis will not perform well when the non-implausible sets are non-convex. The second order specification is given by

$$\begin{aligned} \mathbb{E}[\bar{x}^{\star}_{(i)}(t)] &= \mathbb{E}[x^{\star}_{(i)}(t)], \\ \mathbb{C}ov[\bar{x}^{\star}_{(i)}(t), \bar{x}^{\star}_{(i)}(t')] &= \mathbb{C}ov[x^{\star}_{(i)}(t), x^{\star}_{(i)}(t')] + \delta_{(t,t')}\mathbb{V}ar[\epsilon(t)] \text{ and} \\ \mathbb{C}ov[\bar{x}^{\star}_{(i)}(t), x^{\star}_{(i)}(t')] &= \mathbb{C}ov[x^{\star}_{(i)}(t), x^{\star}_{(i)}(t')]. \end{aligned} \tag{5.2.8}$$

3. Remove all elements of $\chi(t)$ for which $x^{(k)}_{(i)}(t)$ is implausible given the adjusted beliefs about $x^{\star}_{(i)}$. For example, a point is removed if $|Std[x^{(j)}_{(i)}(t) - \mathbb{E}_{\bar{x}^{\star}_{(i)}(T)}[x^{(j)}_{(i)}(t)]]| > 3$

This process is repeated for all p fields in turn. The advantage of this approach is that it is efficient and non-implausible sets are retained.

Why History Matching?

History matching is a tool for statistical inversion; it is a pragmatic, efficient and useful method to tackle large complex problems. It is gaining popularity in the analysis of complex computer models and the methods are beginning to move into other fields. The rise in popularity raises interesting questions about where history matching fits into statistical practice.

History matching is a logical process. For a large number of possible values of the parameter of interest $x^{\star}$, check whether there is evidence that an acceptable match is highly unlikely to be attained. By identifying the set where an acceptable match is highly unlikely to be attained, uncertainty in where the acceptable match is located is resolved. Therefore, history matching is a logical, efficient and understandable way to resolve uncertainty about uncertain quantities. If this set resolves enough uncertainty and the acceptable matches are associated with measurable quantities, it may be used for decision making. If a second

order posterior is required, beliefs about $x^\star$ may be adjusted by the non-implausible set. If a full distributional posterior is required, a full Bayes analysis may be performed within this set making it more efficient [37]. If the set is empty, modelling assumptions or specifications require further consideration.

Suppose the goal of a study was to identify whether a patient does not have osteoporosis. If the non-implausible set of aBMD is identified and no value within the set is indicative of a patient having osteoporosis, the non-implausible set alone is enough to meet the goals of study. Furthermore, suppose the aim is to improve image quality by removing expected scatter, and any non-implausible point provides a good enough match to an X-ray image such that image quality is improved, a non-implausible point is sufficient to achieve the study's goals. In both cases, history matching alone is ample.

A second order specification can be insufficient to describe the complex relationship between quantities. History matching can linearise the relationship and therefore the Bayes linear adjustment will resolve more uncertainty in the parameter of interest. For example, suppose the aim is to calculate the adjusted expectation of a patient's aBMD, the relationship is non-linear and the correlation between i) the summary statistics from the non-implausible set and ii) the system values is larger than that of i) the data and ii) the system values. Therefore, the history matching step enables more uncertainty to be resolved in a Bayes linear analysis.

Given that the prior non-implausible set is consistent with the prior distribution, the posterior non-implausible set is expected to contain the majority of the probability mass of a full Bayes posterior. Therefore, identifying this set narrows down the region from which points are sampled in a full Bayes analysis and will therefore make calculations more efficient. It can also be used to identify if the full Bayes analysis is likely to resolve enough uncertainty to meet the goals of the study. For example, it could identify that certain inputs to a complex computer model are unidentifiable. Therefore, history matching is a fast approximation to the full Bayes analysis which can be used to assess if the study's goals are likely to be met and improve computational efficiency.

History matching leaves open the possibility that all regions of the parameter space are contradicted by the observations and specifications made. When this is the case, it is an obvious indication that the modelling assumptions or belief specifications require further consideration. This is particularly pertinent when the quantities of interest do not have an operational definition and it is questionable whether a true but unknown value

exists. In this case, all that is being asked is if the point is returning an acceptable match to observed behaviour and it is an advantage that the answer can be no.

In summary, history matching is becoming popular because i) it is tractable and computationally efficient, ii) incorporates key uncertainties which have been specified in a logical way and iii) returns an output which can be used for decision making. If decisions cannot be made with the non-implausible set alone, it will inform if the goals of the study are likely to be reached in a full Bayes analysis or resolve more uncertainty in a Bayes linear analysis. If a full Bayes analysis is attempted, it has still been a useful exercise as finding the non-implausible set will make the full Bayes calculations more efficient. Therefore, in many uncertainty analyses, beginning with history matching is a good thing to do.

5.3 History Matching with Random Field Priors

The wave approach to history matching becomes limited for large (over one hundred) input dimensions for at least two reasons.

1. A reasonable number of simulations will only sparsely represent the space. This means that the residual process has almost no effect without strong assumptions on correlation length. The emulator is then effectively a computationally inefficient linear model.

2. Sampling enough points to densely represent the prior space becomes difficult and running the points through every wave impracticable.

Consider the problem of inferring an entity's composition $x_S^\star$ using an X-ray image and G4, see Section 4.3.2. This can be posed as an inverse problem which could be solved by history matching. If the detector has 88×105 pixels, the input dimension is $88 \times 105 \times 2$ and, if the traditional history matching approach is applied, it would not be feasible to check the number of proposal solutions required for an exhaustive search. However, many proposal solutions are improbable given the entities they produce would be physically irrelevant. This implies the conventional specification of the non-implausible set is not exhausting an expert's beliefs. The history matching approach needs to be extended to be able to add in more complex prior beliefs without losing its tractability. In this section, a novel approach to history matching, where prior beliefs are specified using multivariate random fields, will be presented.

When dealing with high dimensional data, the most common trick is finding a lower dimensional set of parameters which describe the majority of the variation. This approach is applied here by adding another layer to the history matching method removing points from the non-implausible set which are physically irrelevant. The system values are assumed to be a set of locations on a multivariate random field. Instead of specifying bounds for $x^\star$, bounds are specified for parameters of a multivariate random field implying a reduced prior non-implausible set.

The general form describing the relationship between simulator and system is given in equation 4.1.1 and 4.1.2 but, for brevity, in this section the relationship between system and simulator is described by equation 5.1.1 where $\epsilon_{md} = 0$. In the next chapter, the case where ϵ_{md} need not equal zero will be considered in more detail. Prior beliefs about $x^\star$ are specified as a multivariate random field of the form

$$x^\star_{(k)}(t) = \mathcal{M}_{(k)}(t) \oplus \mathcal{W}_{(k)}(t) = \sum_m \beta_{(m,k)} h_{(m,k)}(t) \oplus \mathcal{W}_{(k)}(t).$$

In order to perform a history match in this setting, a sample of proposal solutions is required. For β, the same approach as a conventional history match is applied. Plausible bounds are specified and a sample of proposals is produced using a latin-hyper-cube or lattice designs. Then, a large number of samples are made from the residual process using, for example, Gaussian process assumptions (see definition 2.2.5). If Gaussian processes are used to sample the residual process, it may be necessary to either transform or thin the samples so that less probability mass is concentrated near zero. To define the prior non-implausible set, every sample of β is then combined with every sample of $\mathcal{W}(T)$. Once the set is defined, the conventional history matching methodology is applied.

Toy Example: Functional History Match

In order to exemplify this methodology with an infinite dimensional input, a new toy example is required. The example is the convolution of two functions given by

$$\begin{aligned} \mathcal{F}(S) &= (S * g(0), S * g(0.01), \cdots, S * g(1)) \approx \mathcal{F}(x) \text{ where} \\ x &= (S(0), S(0.01) \cdots, S(1)) \text{ and} \\ S * g(t) &= \int_{-\infty}^{\infty} S(\tau) g(t-\tau) d\tau \text{ where} \\ g(t) &= sin(10t). \end{aligned} \tag{5.3.1}$$

$S^\star$ is the function about which an expert would like make inferences and $x^\star$ is a discretisation of $S^\star$. To test the algorithm, $S^\star(t)$ was set as $t + cos(2t)$ and ϵ_{me} was sampled

using a normal distribution with mean zero variance four. Then, the basis functions were specified as $h(t) = (1, t, t^2)$ and the residual process covariance structure was specified using equation 2.2.8 with $l = 0.1$ and $p = 2$, and $\mathbb{V}ar[\mathcal{W}(t)] = 0.01$.

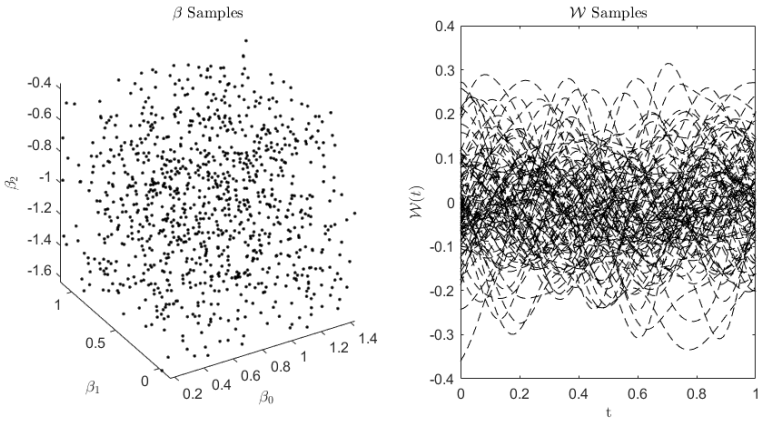

Figure 5.2: Left: samples of β for the toy example using a latin-hyper-cube. Right: samples from $\mathcal{W}(T)$ making Gaussian process assumptions.

Figure 5.2 shows a set of samples of β and $\mathcal{W}(T)$. The full set of proposal solutions is every sampled β with every sampled $\mathcal{W}(T)$ and the implausibility metric used is given by equation 5.2.3. Figure 5.3 shows the output of the history matching procedure. The set of non-implausible functions have a similar shape to the true value and the set of predicted solutions represent the observed behaviour well. After history matching, approximately $1e^{-03}\%$ of the non-implausible set was remaining showing that the procedure can resolve a large amount of uncertainty.

Two Stage History Matching with Random Field Priors

This procedure requires a large number of evaluations of the simulator so, if it is expensive to evaluate, the method becomes intractable. The way in which the prior non-implausible set is defined usually means that the residual process is relatively inactive in that the majority of the variation is described by $\mathcal{M}(t)$. As such, the analysis can be performed in two stages: i) the set of possible regression parameters is narrowed down and ii) the set of possible regression parameters and residual processes are jointly refocused. To ensure this procedure is consistent with the one stage approach, the uncertainty in the output due to the exclusion of the residual process must be taken into account. Otherwise, non-implausible solutions will be incorrectly classified as implausible in the first stage. For stage one, an additional term is added to describe the additional uncertainty due to the

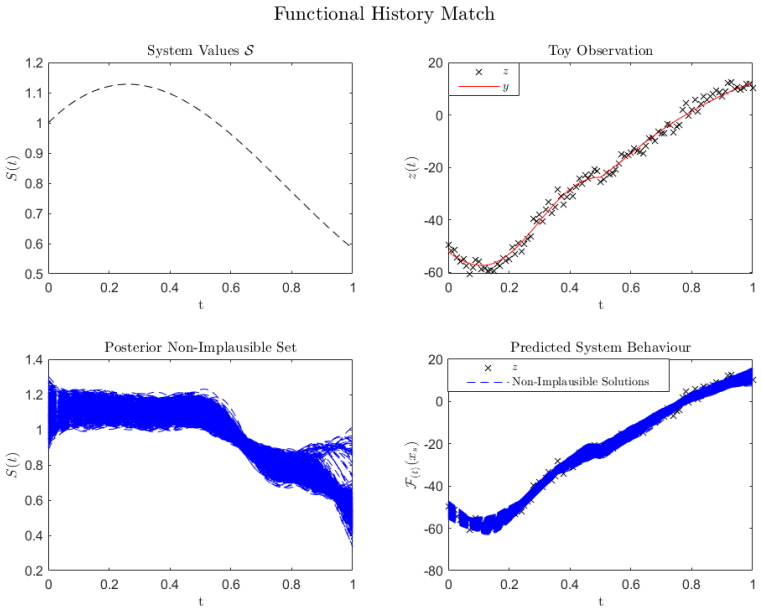

Figure 5.3: Top left: true system value S. Top right: true system behaviour y and sampled observation z. Bottom left: the history matched set of non-implausible functions. Bottom right: predicted system behaviour for all the non-implausible functions.

under specification of the inputs. Then, the implausibility measure is adjusted to account for the additional uncertainty by adding another term to the denominator in, for example, equation 5.2.3. The additional uncertainty may be specified directly or approximated using the following procedure: i) make a small sample of βs, ii) make a small sample of $\mathcal{W}(T)$s, iii) for every sampled regression parameter, evaluate the simulator for every sample of the residual process and calculate the sample variance and iv) make the variance induced by the under-specification the maximum of all those sample variances.

The maximum of the variance is a conservative approach to ensure points are not declared implausible by chance. A more complicated approach where $\mathbb{V}ar[\epsilon_{md}(t)]$ is dependent on t is possible and would be an interesting extension to this work. If $\mathcal{F}$ is expensive to evaluate, it may be replaced by an emulator of the simulator output as a function of β. If β is high dimensional, the first stage may be performed in waves until the size of the set stops changing. This method for history matching with high dimensional input spaces is a simple and tractable extension to history matching methodology. It was designed to keep the advantages of history matching while incorporating a more informed prior specification. The method adds another layer to the history matching procedure to reduce the effective dimension of the space of possible solutions which have to be checked.

Toy Example: Two Stage Functional History Match

Figure 5.4 shows the amount of space cut out during stage one of the two stage analysis. After stage one, 36.1% of the space was remaining meaning that the number of evaluations of the simulator was reduced from 10000^2 to $10000 + 10000 \times 3614$ making the process more efficient.

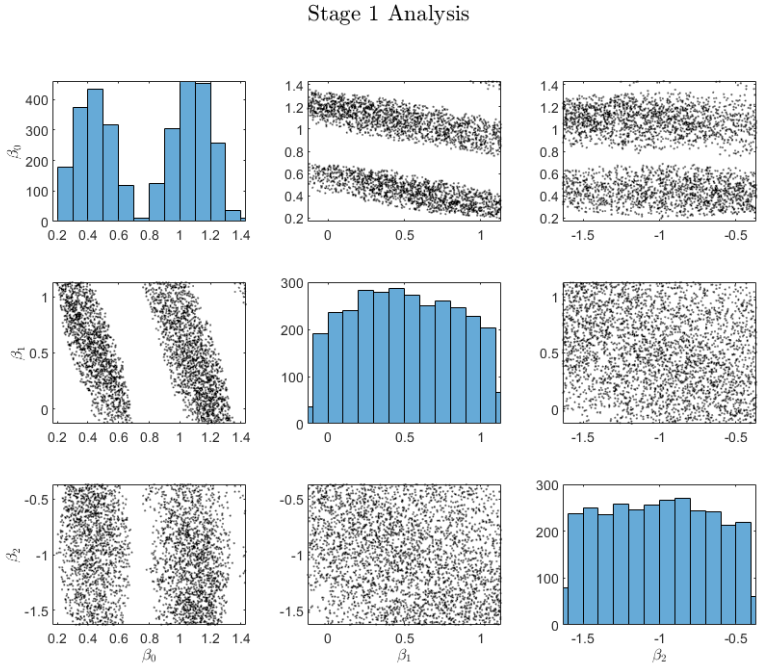

Figure 5.4: Figure of pairs plots, of the non-implausible β after stage one of the two stage functional history match, and marginal histograms.

5.4 Iterative History Matching.

If the emulator is expensive to evaluate, a quicker analysis may be required to provide a single non-implausible point in a useful time-frame. This analysis could be useful for a number of reasons: i) as every non-implausible point represents system behaviour acceptably well, prediction with any point may be enough to provide benefits, ii) if beliefs are adjusted with a summary statistic of a non-implausible set, any point within that set may be good enough to achieve the goals of the study or iii) finding one point can identify an interesting region to begin an exhaustive search. For example, if the aim is to produce an image with the effect of scattered photons reduced and infer entity composition, the clinician can use the scatter corrected image from any non-implausible point while a full

Bayes analysis calculating an expert's posterior distribution for aBMD is performed offline. In this section, a novel iterative approach to finding a set of non-implausible points will be described. The approach is motivated by the use of iterative approaches in signal processing, particularly scatter correction methodologies. The basic idea is that an overestimation of scatter implies an underestimation of thickness which in turn implies an underestimation of scatter. Therefore, it seems reasonable that iterating between estimates will move us to solution with smaller error, which may be non-implausible. It is assumed that the simulator may be re-written in the form

$$\mathcal{F}(x, M = 0) = \mathcal{F}(x, M = 1) + \mathcal{F}(x, M = 2) \tag{5.4.1}$$

where component $\mathcal{F}(x, M = 1)$ has the property that the output of the simulator at location t only depends on the active inputs $x(t)$. Given $\mathcal{F}(x^\star, M = 2)$ is known and the assumption above, inferences about $x^\star$ may be made relatively easily. Given $x^\star$, inferences about $\mathcal{F}(x^\star, M = 2)$ may be made relatively easily. These two observations suggests the following iterative approach.

Algorithm 8 Iterative History Matching

1. Find the adjusted expectation of the system values given the current estimate for $\mathcal{F}(x^{(i)}, M = 2)$

   $$x^{(i+1)} = \mathbb{E}_z[x^\star; \mathcal{F}(x^\star, M = 2) = \mathcal{F}(x^{(i)}, M = 2)]$$

   and calculate implausibility for $x^{(i+1)}$. If enough number of non-implausible points have been found to meet the goals of the study, stop.

2. Otherwise, find the adjusted expectation of the second component given the expected system values

   $$\mathcal{F}(x^{(i+1)}, M = 2) = \mathbb{E}_{D^\mathcal{F}}[\mathcal{F}(x^\star, M = 2); x^\star = x^{(i+1)}]$$

   and return to step 1.

Each of these steps may be individually challenging, however, the number of emulator evaluations required can be reduced to such an extent that searching large spaces becomes feasible in a useful time-frame. It is important to note that iterative searching of the space is a pragmatic decision for computationally challenging problems. The exhaustive

approach presented in the previous section is better in every way apart from being more expensive. Therefore, if inference can be made in a useful time frame using an exhaustive search, then that method should be used.

An important consideration is that, if searching outside the non-implausible set, the error $|\mathcal{F}(x^{(i)}) - z|$ is decreasing through the iterations. If this is not the case, an additional jump control step can be added to help the process find a non-implausible point. There is no guarantee in general that the solution will converge towards the non-implausible set so the method should be used cautiously.

5.4.1 Jump Control

Jump control is the term used in this thesis to ensure an iterative technique is moving in the correct direction. The idea is, given that the most recently checked solution does not have lower implausibility than the previous proposed point, apply a correction to $x^{(i+1)}$ before moving to step 2. The additional step added to Algorithm 8 is

2.5. Otherwise, if $|z - \mathcal{F}(x^{(i+1)})| < \min_i \left\{|z - \mathcal{F}(x^{(i)})|\right\}$, move to step 2. Otherwise, set

$$\mathcal{F}(x^{(i+1)}, M = 2) = \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}(x^{\star}, M = 2); x^{\star} = \mathcal{C}(x^{(i+1)})]$$

where $\mathcal{C}$ is known as a correction function and return to step 1.

Regulating the algorithm in this way can help to ensure that a non-implausible point is found if one exists. There are a number of ways the correction step can be designed which were tested. The first approach is a weighted average of the solutions which have been tried so far

$$\mathcal{C}(x^{(i+1)}) = \sum_{j=1}^{i+1} \alpha_{(j)} x^{(j)} \tag{5.4.2}$$

where the weights $\alpha_{(j)}$, such that $\sum_j \alpha_{(j)} = 1$, are defined by the error between emulator and observation. A second option, which was inspired by iterative signal processing algorithms (Richardson-Lucy deconvolution [90]), is

$$\mathcal{C}_{(j)}(x^{(i+1)}) = A_{(j)} x_{(j)}^{(i+1)}$$

where A is a vector the elements of which are ratios, for example $A = \frac{\mathcal{F}(x^{(i+1)})}{z}$ or $= \frac{z}{\mathcal{F}(x^{(i+1)})}$. This multiplicative approach uses the relative error to dictate how to choose the next solution. The idea is, if the current proposed solution is 50% off the observation, then a good next guess for the system values would be 50% away from the current input.

Although simplistic this algorithm is powerful, however, it is possible to do better.

Given a simulator of the system is available, inferences can be made about what is the expected residual $x^\star - x^{(i+1)}$ given a current residual $z - \mathcal{F}(x^{(i)})$. In this approach, the correction function is of the form

$$\mathcal{C}(x^{(i+1)}) = x^{(i+1)} + \mathcal{M}\left(z - \mathcal{F}(x^{(i+1)}\right) + \mathcal{W}\left(z - \mathcal{F}(x^{(i+1)}\right).$$

Beliefs about $\mathcal{M}, \mathcal{W}$ are adjusted using the simulator. A set of samples $\{x^{(i)}\}_{i=1:\mathbb{N}}$ are taken from the prior non-implausible set. Then, for every sample, the simulator is evaluated giving the set $\{\mathcal{F}(x^{(i)})\}_{i=1:\mathbb{N}}$. The $\mathbb{N}^2$ differences between every pair of simulator evaluations $\{\mathcal{F}(x^{(i)}) - \mathcal{F}(x^{(j)})\}_{i=1:\mathbb{N},j=1:\mathbb{N}}$ are then used to adjust beliefs about the correction function. The relationship between correction functions and simulator evaluations is

$$x^{(i)} - x^{(j)} = \mathcal{M}\left(\mathcal{F}(x^{(i)}) - \mathcal{F}(x^{(j)}) + \epsilon^{(i,j)}\right) + \mathcal{W}\left(\mathcal{F}(x^{(i)}) - \mathcal{F}(x^{(j)}) + \epsilon^{(i,j)}\right)$$

where $\epsilon^{(i,j)}$ is a sample from the measurement error distribution. The method effectively uses the simulator to adjust beliefs about the relationship between the residual from i) expected to observed system behaviour and ii) proposed to actual system value. The correction function can be difficult to fit and a large number of simulator evaluations can be required to fit it. If efficient real-time inference is required, then the offline cost can be acceptable. The approach that is used in the examples in this thesis was 5.4.2. This was because the multivariate random field nature of the input caused Richardson-Lucy to fail and the dimension of the input was too large for a simulator trained correction factor. However, for the toy example, both Richardson-Lucy and simulator trained correction performed better so, for lower dimensional problems, these approaches can perform well.

5.5 History Matching for Alloy Identification with a MAP

History matching is usually applied to find a set of inputs to a complex computer model which are acceptable matches to observed system behaviour. In this section, history matching will be applied to an uncertain look-up problem. The example is from the X-ray industry and the task is to infer the thickness and alloy of a material using a MAP equipped detector. The method is to use observations with known system values to adjust beliefs about the forward operator and history matching is used to make inferences about entities with unknown system values, see Section 5.2.1 subsection “History Matching for

Uncertain Look-ups". For this example, it is assumed that the i) scatter-to-primary-ratio (see definition 3.5.1) is small enough that it may be ignored and ii) forward operator is independent of the position on the detector.

Inferring thickness and alloy of an entity is of interest as it can be transformed into aBMD. If a MAP equipped detector can return statistics informative for aBMD from a single exposure, these statistics may be calculated for any exposure on a MAP equipped detector. This reduces the need for referrals and increases the chance of diagnosing a symptom-less patient undergoing routine treatment. Therefore, it is interesting to show that MAP technology can be used to infer the alloy and thickness of an entity.

Given an observation of an entity, an expert would like to infer the entity composition in the square pyramid (SP) with base on a pixel and point on the source. The observation process is given by $z = (\beta_0^{EC}, \beta_1^{EC}) = y \oplus \epsilon_{me}$ where $(\beta_0^{EC}, \beta_1^{EC})$ are estimates for the continuous energy parameters given in equation 3.5.4. The problem set up is as follows: i) the thickness and alloy are notated using the two-vector $x_S = (\mu, \eta)$, ii) the true sample composition is notated $x_S^\star$ and iii) the relationship between $x_S^\star$ and y is modelled via a functional relationship

$$\mathcal{F}_{(k)}(x_S^\star) = y_{(k)} = \sum_{m=1}^{6} \beta_{(m,k)} h_{(m)}(x_S^\star) \oplus \mathcal{W}_{(k)}(x_S^\star).$$

The basis functions were specified as $h(x_S) = (1, \mu, \eta, \mu\eta, \mu^2, \eta^2)$ and the residual process covariance structure was specified as separable (equation 2.2.7) with correlation function given by equation 2.2.8 with $l = (0.2, 2)$ and $p = 2$. The set of samples is notated $\mathcal{S} = \{S^{(i)}\}_{i=1:N}$ of which observations $D^f = \{z^{(i)}\}_{i=1:N}$ are made. Each $S^{(i)}$ has a known alloy and thickness, notated $X = \{x_S^{\star(i)}\}_{i=1:N}$. The training samples X are chosen so that the range of possible values a new observation can take are contained within the convex hull of X. This means that problems associated with forecasting outside that convex hull need not be considered. The prior non-implausible set is then specified as the convex hull of X. For each training observation, the model is given by equation 5.2.6. The measurement errors are assumed to be expectation zero, constant variance and mutually uncorrelated. To perform the history match, the implausibility measure was the maximum implausibility (equation 5.2.3) with a threshold of 3. The adjusted expectations were approximated using the GLS approach, see key-formula 2.2.1. The inference is validated by cross validation where i) every observation is excluded from the training-set, ii) beliefs about $\mathcal{F}$ adjusted and iii) the non-implausible set for the excluded observation found. The first stage of

the algorithm is considered valid if the true $x_S^{\star(i)}$ is in the associated non-implausible set $\chi^{(i)}$ more than 95% of the time and the posterior non-implausible set is smaller than the prior. The approach was tested in simulation using a simulation proxy for a MAP equipped detector. Two pencil beam simulations were performed for every training entity each with two distinct known thicknesses of thin nickel directly in front of the detector. Two further simulations were performed without the entity but with the nickel filtration to act as I_0, see equation 3.5.2. These simulations were used to calculate estimates for the continuous energy parameters, see equation 3.5.4. Only the central pixels of the pencil beam simulations were used so the effect of scatter is minimised.

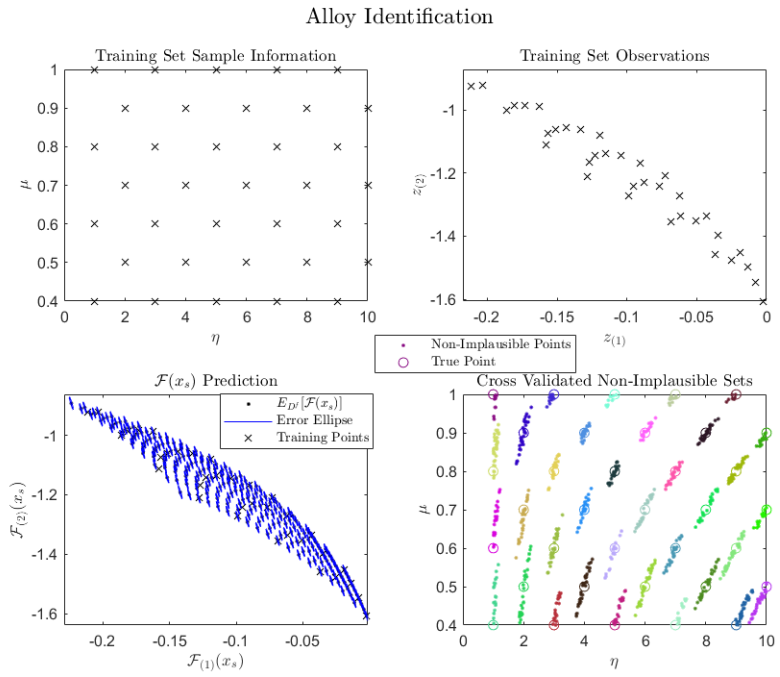

Figure 5.5: Top left: training system values. Top right: simulated continuous energy parameters. Bottom left: emulator output. Bottom right: cross validated non-implausible sets.

Figure 5.5 illustrates the algorithm and the cross validation procedure. Top left shows x_S for the training simulations. Top right shows the associated continuous energy parameters calculated from the G4 simulations. Bottom left shows the adjust beliefs about $\mathcal{F}(x_S)$. Bottom right shows the cross validated non-implausible sets for every entity in the training-set. The sets have been significantly reduced and the known truth is always in the non-implausible set which is why the method was considered valid.

The outputs of the cross validation procedure were then used make a second order specification for $\{x_S^\star, \bar{x}_s^\star\}$. Figure 5.6 illustrates this second order specification. The cor-

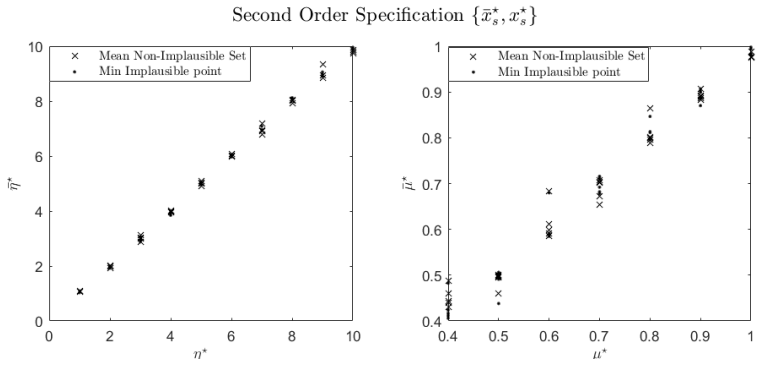

Figure 5.6: Left: relationship between thickness and summary statistics from the non-implausible sets. Right: relationship between material and summary statistics from the non-implausible sets.

relation between system value and summary statistics of the non-implausible set is larger than the system value and the observation, therefore, the Bayes linear adjustment with $\bar{x}^\star$ will resolve more uncertainty than the adjustment with z.

Discussion: History Matching for Alloy Identification with a MAP

In this section, a method for inferring the composition of an entity using observations with a MAP equipped detector has been presented. The method was shown in simulation assuming there was no scatter or spatial error on the detector. Given these assumptions are correct, this work provides a proof-of-concept that inferring an entity's composition is possible using a single MAP equipped flat panel exposure. The problem is that the error induced by scatter and spatial variation means that the uncertainty resolution in the real world is too small to be useful. This was found by applying this method using a system equipped with and without an ASG. When there was no ASG, the results were poor and the ASG did not reduce scatter enough to make the inferences returned useful. Essentially there was still large entity dependent spatial variation across the image which could not easily be resolved. The initial idea in this study was to correct scatter and use a method using a MAP to infer aBMD; this section has presented the method we devised. We found that in order to correct scatter effectively, the thickness and alloy of the entity must be inferred. Given this, it is more appropriate to history match G4 to provide both scatter correction and composition inference. Presenting this approach will be the focus of the following section.

5.6 History Matching GEANT4 (G4) for Scatter Correction

In the previous chapter, our method for emulating G4 − in other words predict scatter given an entity's composition − was presented. In this chapter, we will make inferences about an entity's composition from an X-ray image simulated using a divergent beam evaluation of G4. The problem of inferring entity composition will be posed as an inverse problem where beliefs about the system values are organised as a mixture of multivariate random fields. This study will illustrate the method and serve as an initial proof point that, from a mono-spectral digital radiograph, Bayesian inverse problem solving can be used to return useful inferences about an entity's composition. Furthermore, the inferred entity composition can be used to reduce the impact of scattered photons resulting in improved image quality. As the proof is in simulation, the proof point is not that the method will work in the real world but, if this point could not be proved, there is strong evidence the method would never work in the real world. How to transfer this method to the real world is considered in the following chapter.

The inverse problem set up is

$$\begin{aligned} \mathcal{F}(S^{G4}) &= y \oplus \epsilon_{me} \text{ where} \\ y &= \mathcal{F}'(x_S^\star, M=0) \oplus \epsilon_{md} \text{ and} \\ \mathcal{F}'(x_S^\star, M=0) &= \mathcal{F}'(x_S^\star, M=1) + \mathcal{F}'(x_S^\star, M=2). \end{aligned}$$

As G4 is stochastic, y is G4's population mean and ϵ_{me} is the residual from that mean. ϵ_{md} is the difference between the population mean of G4 and the hierarchical-superposition emulator. Therefore − and for the reasons discussed in sections 4.3.5 and 4.4 − the emulator acts as the *de-facto* simulator and a simulation acts as the observation. $\mathcal{F}'(x_S^\star, M=1)$ is defined by equation 4.3.6 and $\mathcal{F}'(x_S^\star, M=2)$ by equation 4.3.10. The Direct-Beam-Emulator was calculated directly while the scattering component utilised look-up-tables to speed up calculations.

The discretisation $x_S^\star = \mathcal{D}(S^{G4})$ is defined by equation 4.3.1. The implausibility measure used is the mean of the pixel wise implausibility. This was chosen because the output vector had thousands of elements so a max implausibility measure is likely to declare the system value implausible by chance. When using mean implausibility, there are a couple of important features which have to be considered. The averaging means that an image where every pixel has implausibility 3 can be non-implausible but one which has very small implausibility in all but one pixel can be implausible. There is nothing

incorporated in this algorithm to test for these features. The approach we have taken is to produce diagnostic images which are returned to the user. The implausibility image and the percentage error image are returned and, if there are large areas which are not well modelled, it can be diagnosed. The functional history match will be tested on a simple entity and why it is not used for more complex entities explained. Finally, the iterative approach will be applied to a more complex entity. This is the algorithm which will be applied in the real world in the next chapter. The algorithms present the general algorithm while the actual numbers are used for the example in this section.

The resolution of the real world detectors of interest had approximately 9000000 pixels. Solving a problem with this resolution is challenging for a number of reasons: i) the number of points in the emulator training-set would be huge, ii) the simulator output is much steeper making it harder to emulate and iii) searching the space would be impractical without further methodological development. The adjusted beliefs about entity composition are still useful if it is not high resolution. To be useful, the expected scatter is required at the resolution of the original detector. However, scatter is smooth (in other words only contains low frequency components) so a simple up-sampling using B-splines is enough to provide image quality improvements. Therefore, i) the simulator detector was changed to mimic a real world detector, ii) the output is sub-sampled, iii) the algorithm applied to the sub-sampled image and vi) the scatter estimate is up-sampled to produce expected scatter of required resolution. The model discrepancy and measurement error variance were specified as

$$\mathbb{V}ar[\epsilon_{md} + \epsilon_{me}] = 0.05(\mathcal{SS}(y^2) - \mathcal{SS}(y)^2) \tag{5.6.1}$$

where $\mathcal{SS}(y)$ is a sub-sampling function which returns the average of a region within a square grid. This was chosen because G4's stochastic component is approximately Gaussian and pixel wise independent. Therefore, the uncertainty in the population mean (population mean estimator error) of the sub-sampled regions was considered appropriate to describe the measurement error after sub-sampling. It was assumed that the model discrepancy was proportional to $(\mathcal{SS}(y^2) - \mathcal{SS}(y)^2)$ and 0.05 was chosen by checking that the implausibility for the validation samples was not too small (mean implausibility was around 1.5). This model discrepancy specification is only used in this example for exemplifying the methodology in simulation. When the methods are transferred to the real world, a more careful approach is utilised (see Chapter 6).

Prior Specifications Sample One

A cuboid block of constant material and thickness 6cm is used to exemplify the exhaustive search method. The material was an alloy of PMMA (80%) and aluminium (20%). The range of (i, j) in the region illuminated by a five degree divergent beam was normalised to $[0, 1]$ to make the prior specification easier. The prior specification for the thickness field is

$$\eta^\star(i, j) = (1, i, j, i^2, j^2)\beta^\eta \oplus \mathcal{W}^\eta(i, j).$$

To create a proposal solution, the parameters β^η were sampled from a latin-hyper-cube in the range $[4, -0.5, -0.5, -1, -1] \to [9, 0.5, 0.5, 1, 1]$. This prior specification was made by considering sampled outputs of the random field. For example, the thickness could range between 4 and 8 and the sample could be flat to non-convex with a turning point not too close to the edge. Asking an expert to specify a regression parameter can be unreasonable but asking them to specify what the field would look like and eliciting beliefs about regression parameters based on that is more reasonable. The residual process had variance 0.005 and correlation structure specified by equation 2.2.8 with $l = 5$ and $p = 2$. To get a proposal solution, samples from the residual process were made using Gaussian process assumptions, see definition 2.2.5. The material specification was assumed to be a constant alloy of PMMA and aluminium i.e. $\mu^\star(i, j) = \beta_0^\mu$ where proposals for β_0^μ were sampled uniformly on $[0.6, 1]$; 1 being 100% PMMA. This specification is encoding the belief that the entity's thickness field is smooth and the entire entity is made of the same material. The history match with random field priors was performed using the two stage analysis described in Section 5.3. $\{\beta^\eta, \beta_0^\mu\}$ were focused first to narrow down the space, and then $\{\beta, \beta_0^\mu, \mathcal{W}(T)\}$ were jointly refocused.

Sample One Results

Figure 5.7 left shows the standardised error $Std(\eta^\star(i, j) - \bar{\eta}^\star(i, j))$ where $\eta^\star(i, j)$ is assumed to be in the form given in equation 5.2.7. Here $\bar{\eta}^\star(i, j)$ is the mean summary statistic (see equation 5.2.5). As the value is never over three, there is no evidence to suggest the known truth is not the non-implausible set which does not invalidate the process. The right plot shows the prior and posterior non-implausible sets for μ. 91% of the variance has been resolved and large portions of the space have been ruled out completely.

Figure 5.8 shows the uncertainty resolution in the thickness field. The top plot shows the prior and posterior expected field with associated two sigma credible interval. The

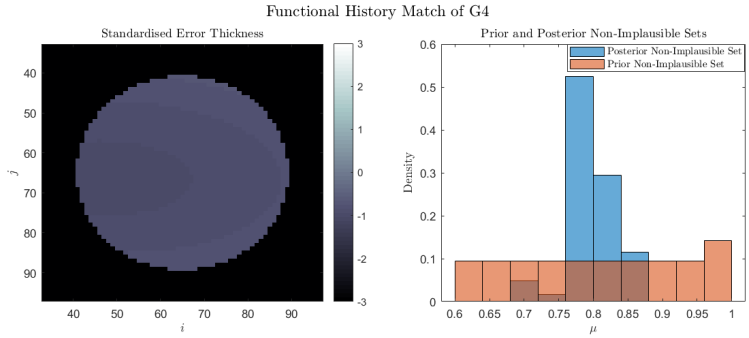

Figure 5.7: Left: standardised error $Std(\eta(t) - mean(\chi^\star(t)))$. Right: prior and posterior non-implausible sets for μ.

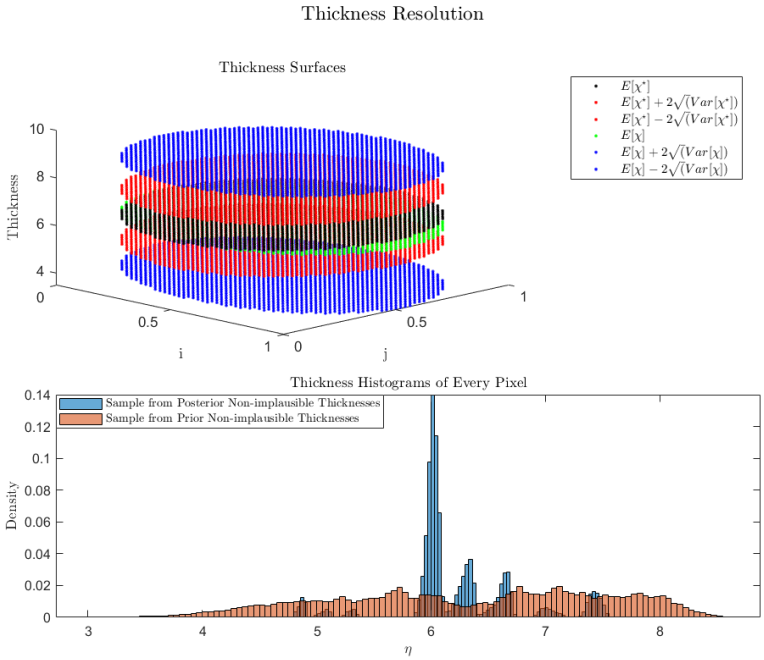

Figure 5.8: Top: second order summary of the non-implausible thickness fields. Bottom: non-implausible thickness field histograms.

known truth is within two sigma of the posterior expectation which does not invalidate the method. The bottom plot shows the histogram of the prior and posterior non-implausible sets for every location on the fields. A large amount of the variation has been resolved, however, the posterior is multi-modal which should be considered if a Bayes Linear adjustment using a summary statistic is attempted.

This example is an important proof point. It shows, given G4 is the real world, that the problem is not in the Bayesian sense ill-posed and that a large amount of uncertainty, in both thickness and alloy, can be resolved. This implies that diagnostic measures, like

aBMD, may be inferred from mono-spectral digital radiography systems. Furthermore, even if the entity's composition is uncertain, the effect of scatter may be inferred and removed from an image. This is not to say that the method will transfer to the real world and more complex samples. However, if this test could not be passed, there would be strong evidence to suggest that the method could never work.

The plan was to use an emulator of G4 to history match a patient's composition. There were two problems which we struggled to overcome. The first is that the emulator built of G4 was still slow. This meant checking the required number of solutions was not possible in a useful time-frame. An option was to build an emulator of our emulator of G4 as a function of the regression parameters of the system value's random field. This would speed up the process but was limited by the amount of time taken to refocus the system value's residual process. Another option was to use a more flexible set of basis functions with a larger number of parameters and remove the residual process. The problem with this approach was that the number of parameters required to describe complex anatomy like a NOF would be large and it would be difficult to make a wave based history matching approach efficient and robust. A further option would be to combine the suggestions above with sequential history-matching of small subsets of the output image. However, given that for every output location 9000 ray traces and emulator evaluations were required, this was still too expensive. These considerations led us to develop the iterative approach which will now be described. This approach enables non-implausible points to be found with under 40 evaluations of the emulator of G4.

5.6.1 Iterative History Matching for Scatter Correction

The pivotal property enabling the iterative approach to history matching G4 is

$$\mathcal{F}'_{(i,j)}(x_S^\star, M = 1) \perp\!\!\!\perp x_S^\star | x_S^\star(i,j).$$

This is specifying that the output of the direct beam simulator (see equation 4.3.3), at location (i,j), is conditionally independent of $x_S^\star$ given $x_S^\star(i,j)$. Therefore, given that the scattering component is known, the inverse problem is relatively easy to solve as it can be split up into a set of two dimensional inverse problems − one for every pixel. This implies that an iterative technique may provide an efficient method to find a non-implausible point. The iterative process is as follows.

1. Prediction step: given a proposed solution $x_S^{(i)}$ for $x_S^\star$, calculate

$$\mathbb{E}_{D^\mathcal{F}}[y; x_S^\star = x_S^{(i)}]$$

and calculate the implausibility measure. If non-implausible, stop. Otherwise, go to step two.

2. Jump control step: if $\mathcal{IM}(x_S^{(i)}, z) < \min\{\mathcal{IM}(x_S^{(i-1)}, z), \cdots, \mathcal{IM}(x_S^{(1)}, z)\}$ go to step 3. Otherwise, set $x_S^{(i+1)} = 0.5x_S^{(i)} + 0.5x_S^{(j)}$ and return to step one. Here j is the minimum non-implausible point in the set with an average error of opposite sign. In other words, if the average error

$$\frac{1}{\theta\kappa}\sum_t \left(z - \mathbb{E}_{D^\mathcal{F}}[y_{(t)}; x_S^\star = x_S^{(i)}]\right) > 0,$$

then $j = \text{argmin}_{k\in\mathcal{K}}\{\mathcal{IM}(x_S^{(k)}, z)\}$ where $\mathcal{K}$ is the set of ks such that

$$\frac{1}{\theta\kappa}\sum_t \left(z - \mathbb{E}_{D^\mathcal{F}}[y_{(t)}; x_S^\star = x_S^{(k)}]\right) < 0.$$

3. History matching step: find the non-implausible set for $x_S^\star$ using the relations

$$y - \mathbb{E}_{D^\mathcal{F}}[\mathcal{F}(x_S^{(i)}, M = 2)] = \mathcal{F}(x_S^\star, M = 1) \text{ and}$$
$$\mathcal{F}(S^{G4}) = y \oplus \epsilon_{me}.$$

This is perform in two stages: i) calculate the set of non-implausible sets $\chi^{(i)} = \{\chi^{(i)}(t^{(j)})\}_{j=1:\kappa\theta}$ for every pixel and ii) given these non-implausible sets and beliefs specified about $x_S^\star$, calculate the approximate adjusted expectation

$$\mathbb{E}_{x_{\bar{S}}^\star(T)}[x_S^\star; \mathcal{F}(x_S^\star, M = 2) = \mathbb{E}_{D^\mathcal{F}}[\mathcal{F}(x_S^{(i)}, M = 2)]]$$

using the method outlined in Section 5.2.1. This approximate adjusted expectation is used as the next proposal solution $x^{(i+1)}$.

This iterative procedure was decided on after trying a number of methods. The initial attempt used no jump control but it could get stuck jumping between two implausible solutions. The Richardson-Lucy type approach did not help matters as it overcompensated both thickness and material fields together often finding a worse solution. The emulator trained correction factor was too difficult to fit as the dimension of the inputs was too large. The proposed approach has worked well for a number of entities consistently finding a non-implausible point. The method described above is summarised in Algorithm 9 and

is used for a number of examples in the following two chapters. In the rest of this thesis, Algorithm 7 will take x'_S (see Algorithm 7) as an input and proceed from step 2. Each of the steps will now be discussed in detail and algorithms stated when necessary.

Algorithm 9 Iterative G4 History Match

Inputs: z, $\mathcal{LU}$, Direct-Beam-Emulator, ImThresh

Outputs: $\hat{x_S}$, $\chi_s^\star$

1: Specify initial proposal solution $x_S^{(1)}$.

2: **for** i in $1, 2, \cdots$ **do**

3: Prediction step: using $x_S^{(i)}$ calculate the expected system behaviour

$$\mathbb{E}_{D^{\mathcal{F}}}[y; x_S^\star = x_S^{(i)}]$$

using Algorithm 7.

4: Calculate implausibility $\mathcal{IM}(x_S^{(i)}, z)$.

5: **if** $\mathcal{IM}(x_S^{(i)}, z) <$ ImThresh **then**

6: Set $\hat{x_S} = x_S^{(i)}$

7: **break loop**

8: **end if**

9: **begin:** Jump control step.

10: **if** $\mathcal{IM}(x_S^{(i)}, z) < \min\{\mathcal{IM}(x_S^{(i-1)}, z), \cdots, \mathcal{IM}(x_S^{(1)}, z)\}$ **then**

11: Set $x_S^{(i+1)} = 0.5x_S^{(i)} + 0.5x_S^{(j)}$ where j is the minimum non-implausible point in the set with an average error of opposite sign.

12: **end**

13: **else**

14: History matching step: set $x_S^{(i+1)}$ and $\chi_s^\star$ to the outputs of Algorithm 13 which makes inferences about the alloy and thickness of an entity given the direct beam image.

15: **end if**

16: **end for**

Prior Specifications Entity Two

To test the algorithm, an entity designed to be analogous to an arm was used. It is an annulus made of PMMA with the centre filled with a cylinder of aluminium. The PMMA annulus had radius 4cm and the aluminium cylinder 1cm. A diagram is shown

in Figure 4.14 with $r_1 = 1$ and $r_2 = 8$. This entity geometry is complex enough that an isotropic approach to modelling G4 would incur a large error, making it a good test of the hierarchical-superposition approach presented in this thesis.

The prior beliefs about $x_S^\star$ were specified as a mixture of multivariate random fields, given by

$$\eta^\star(i,j) = (1, i, j, ij, i^2, j^2)\beta^\eta \oplus \mathcal{W}^\eta(i,j) \text{ and}$$
$$\mu^\star(i,j) = \sum_{\omega=1}^{\tilde{q}} \pi_{(\omega)}(i,j)\beta^\mu_{(:,\omega)} \oplus \mathcal{W}^\mu_{(\omega)}(i,j).$$

Here $\pi_{(\omega)}$ is an indicator field defining a partition. The prior segmentation of the field is shown in Figure 5.9.

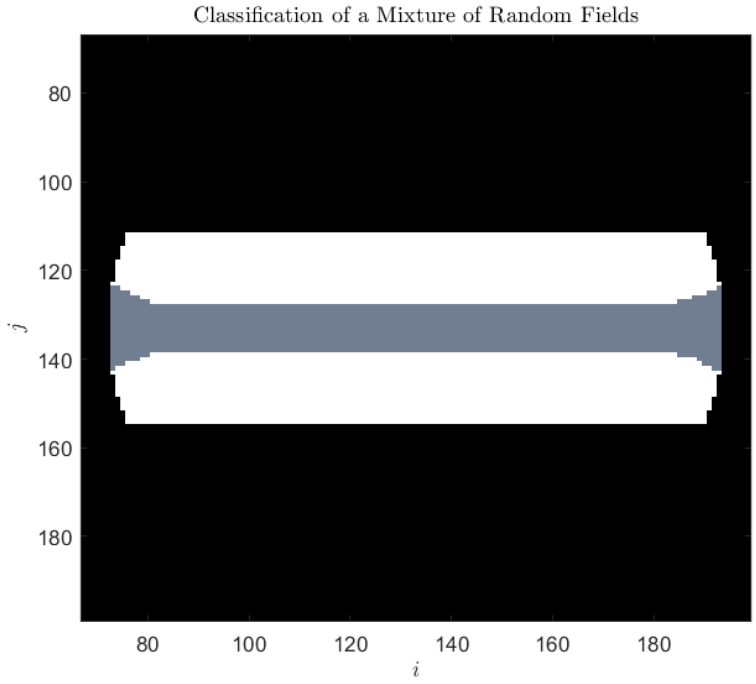

Figure 5.9: Prior segmentation for the material field. White - segment one (known material), grey - segment two, black - no entity (open beam).

The second order specification for the first segment, white in figure 5.9, was given by

$$\mathbb{E}[\beta^\mu_{(1,1)}] = 1,\ \mathbb{V}ar[\beta^\mu_{(:,1)}] = 0 \text{ and } \mathcal{W}^\mu_{(1)} = 0.$$

As tissue regions can be classified using i) a threshold, ii) image processing techniques or iii) user interaction, it was considered reasonable that a region of a clinical X-ray image could be identified as known 100% tissue *a-priori*. The proxy in this example is assuming

that a region can be identified as known 100% PMMA. Assuming a region can be known *a-priori* makes the iterative process more efficient and stable. Other approaches without this assumption were tested but they required five times the emulator evaluations. For the second segment, the sample mean of the non-implausible set was used as a point estimate for the adjusted expectation. This is essentially specifying that the pixels in that segment are uncorrelated with one another. This was chosen to represent a weak prior about the structure of the material field where it is not known to be PMMA. As this approach is used for the specification of many entities in this thesis, the method is summarised for a general entity in Algorithm 10. A more complex prior specification is possible and would be an interesting extension to the method. For example, the thickness could be specified as a mixture with more components and B-splines could be used to specify beliefs about the mean function, see Section 2.3. However, for the purposes of this study, the form in Algorithm 10 was enough to improve image quality and return useful diagnostic measures.

Algorithm 10 Prior Specification of Thickness and Alloy fields.

Inputs:

Outputs: Prior-Specification-for-$x_S^\star$

1: Specify the region which is *a-priori* known as material one.

2: Specify the basis functions for thickness field.

3: Specify correlation length and variance of thickness field's residual process.

History Matching Step

The first procedure applied in the history matching step is to find the non-implausible set for every pixel on the image. Bounds are set for possible thicknesses and alloys, and a set of values χ_s is sampled from the implied set using a lattice. For this stage, the implausibility measure was given by equation 5.2.1 with a threshold of three. The method for finding a non-implausible point for a single pixel is summarised in Algorithm 11. Figure 5.10 left shows the Direct-Beam-Emulator output for the prior and posterior non-implausible sets for one pixel in partition two. All non-implausible points are within the error bars showing that the process is working as expected. The right hand plots show the prior and posterior sets, there is clear positive correlation in the posterior set and it is non-convex. The method for finding the non-implausible set for every location on the field is

Algorithm 11 History Match Thickness and Alloy for Single Pixel

Inputs: Direct-Beam-Emulator, $\tilde{z} = y(i,j) - \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}_{(i,j)}(x_S^{(j)}, M = 2)]$, χ_s, ImThresh

Outputs: $\chi_s^\star(i,j)$

1: Set $\chi_s^\star(i,j) = \chi_s$.
2: **for** i in 1 : nrow(χ_s) **do**
3: Using the Direct-Beam-Emulator and $\chi_{s(i)}$, calculate implausibility $\mathcal{IM}_{(i)} = \mathcal{IM}(\tilde{z}, \chi_{s(i)})$.
4: **if** $\mathcal{IM}_{(i)} >$ ImThresh **then**
5: Remove $\chi_{s(i)}$ from $\chi_s^\star(i,j)$.
6: **end if**
7: **end for**

Algorithm 12 History Match Thickness and Alloy for Every Pixel

Inputs: Direct-Beam-Emulator, $\tilde{z} = y - \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}(x_S^{(j)}, M = 2)]$, χ_s, ImThresh

Outputs: $\chi_s^\star(T)$

1: **for** i in 1:nrow(T) **do**
2: Using Direct-Beam-Emulator, χ_s and $\tilde{z}(t^{(i)})$, apply Algorithm 11 to calculate $\chi_s^\star(t^{(i)})$.
3: **end for**

summarised in Algorithm 12. Once the non-implausible sets are calculated, the next step is to adjust beliefs about the multivariate random fields given these sets. The approach uses the mean summary statistic (equation 5.2.5) at every location where the material is known to adjust beliefs at every location in T. The more obvious approach would be to use the summary statistics at every location to adjust beliefs. However, the non-convex nature of the posterior non-implausible sets means that the summary statistic was itself non-implausible. This caused the adjusted thickness alloy combination to be implausible, even if the expected scatter was correct, and the iterative approach to fail. If the alloy is known, the mean of the non-implausible set is itself non-implausible and the problem does not occur. In other words, if the alloy is known, the set is convex.

1. The approach begins by calculating the mean and variance of the thickness non-implausible sets where the material is known. Figure 5.11 shows the mean summary statistic $\bar{\eta}^\star(T')$ with $2\sqrt{\sigma_\eta(T')}$ error bars where T' is the set of locations on the field in partition one and $\sigma_\eta(T')$ is the variance of the non-implausible set. This

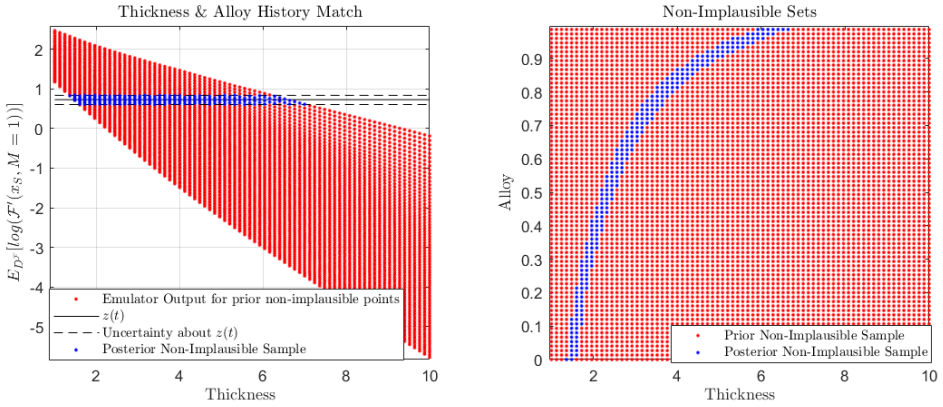

Figure 5.10: Left: Direct-Beam-Emulator output for implausible and non-implausible sets. Right: input space partitioned into implausible and non-implausible sets.

illustrates the set of locations on the field used to adjust beliefs about $\eta^\star$. The variance is relatively small as, given $\mu^\star(i,j)$ is known, the set of possible thickness is small.

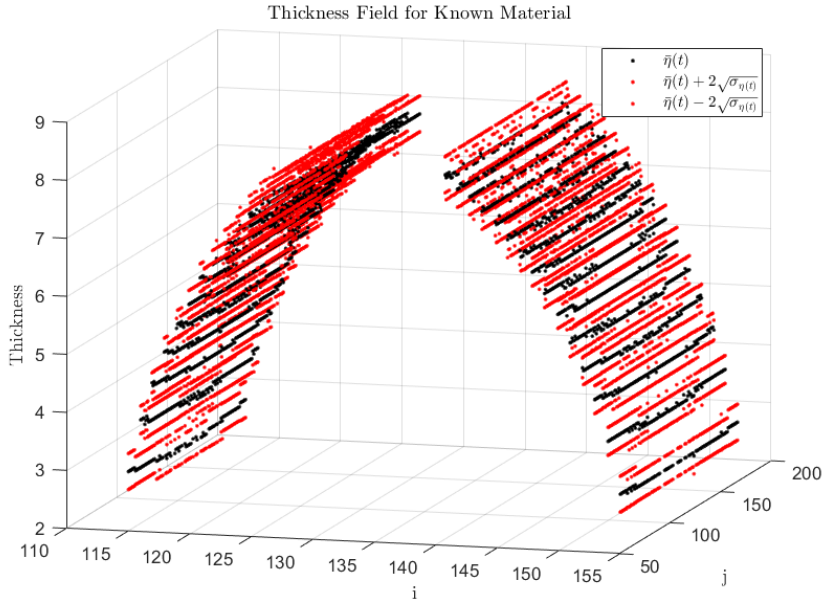

Figure 5.11: Second order statistics for the partition where the material is known a-priori.

2. Then to adjust the thickness field at all locations in T, it is assumed that the summary statistic is related to system values via equation 5.2.7. The second order specification is given by equation 5.2.8 and approximate adjusted beliefs are calculated using equation 2.1.3. Figure 5.12 shows the adjusted beliefs about the thickness field given the non-implausible sets for the partition with known material.

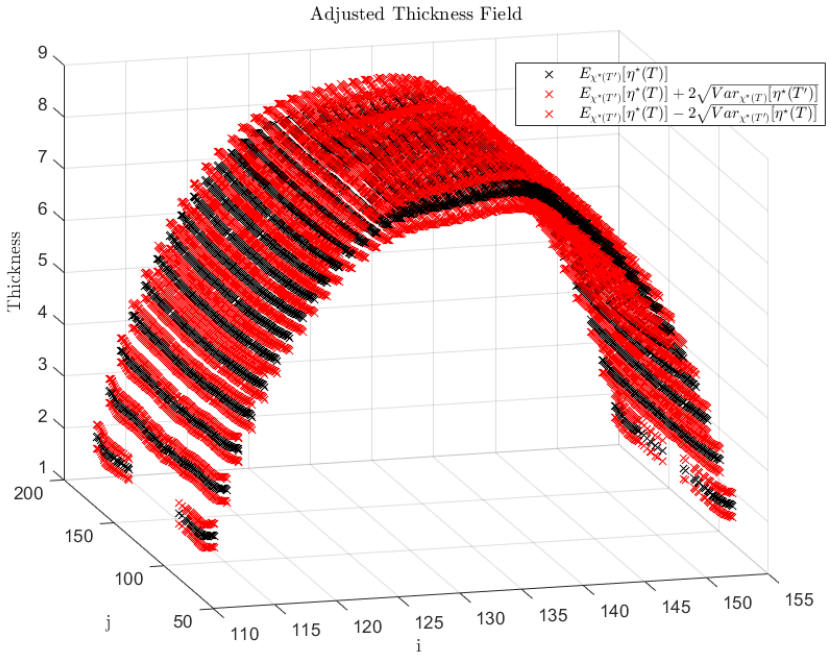

Figure 5.12: Adjusted beliefs about thickness field given non-implausible sets for the partition where the material is known.

3. Given the adjusted beliefs about the thickness field, the non-implausible sets for the partition where the alloy is unknown are refocused. This involves removing all values that have implausibility greater than three given the adjusted beliefs about the thickness field. Finally, the sample mean of the refocused non-implausible sets are used as the point to check in the next iteration. Figure 5.13 illustrates trimming the non-implausible sets given the adjusted beliefs about the thickness field. The size of the set has been significantly reduced showing that a large amount of uncertainty in the entity composition may be resolved with this method. Figure 5.14 shows an image of the adjusted expectation of the alloy field given the refocused non-implausible sets. This result is then passed along with the adjusted thickness field to the next iteration.

Given that expected scatter has been calculated, the summary of the history matching step is given by Algorithm 13. Then the process of history matching G4 – finding a non-implausible point $\hat{x_S}$ and the non-implausible sets $\chi_s^\star$ given that the scatter component is calculated using $\hat{x_S}$ – is given by Algorithm 14.

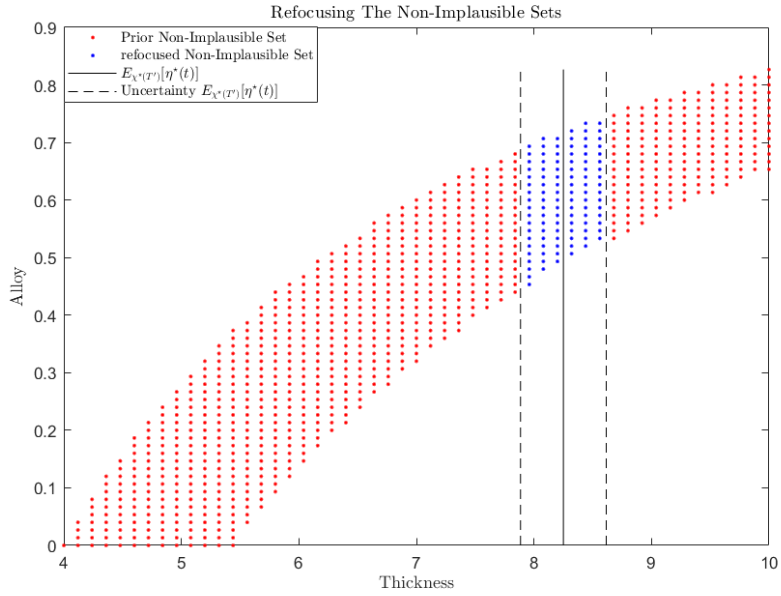

Figure 5.13: Refocused non-implausible set for a single pixel where the alloy is uncertain (partition two).

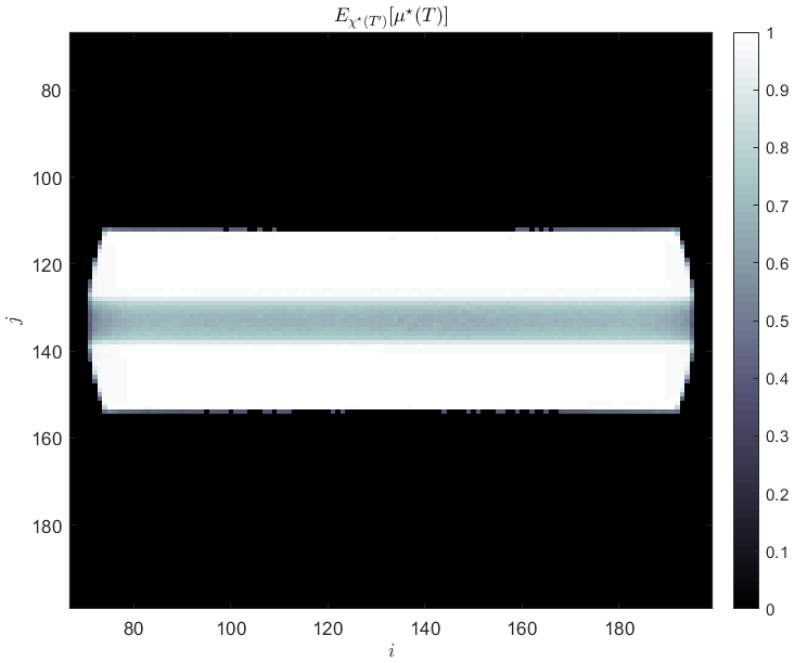

Figure 5.14: Estimate of the alloy field for one iteration which is passed to next iteration.

Iterative History Matching Results

This subsection will present the results of the iterative approach to history matching G4 in simulation. Generally the first proposal solution would be the expected system value

Algorithm 13 Adjust beliefs about Thickness and Alloy Fields

Inputs: Direct-Beam-Emulator, $\tilde{z} = y - \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}(x_S^{(j)}, M = 2)]$, χ_s, ImThresh, Prior-Specification-for-$x_S^\star$

Outputs: $\hat{x_S}$, $\chi_s^\star$

Using Direct-Beam-Emulator, $\tilde{z} = y - \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}(x_S^{(j)}, M = 2)]$, χ_s and ImThresh, perform Algorithm 12.
for i in 1:nrow(T') **do**
    Calculate the sample mean and variance of $\chi_s^\star(t^{(i)})$.
end for
Using Prior-Specification-for-$x_S^\star$ and $\chi_s^\star(T')$, calculate the adjusted expectation and variance of the thickness field everywhere $\mathbb{E}_{\chi_s(T')}[\eta(T)]$ and $\mathbb{V}ar_{\chi_s(T')}[\eta(T)]$.
Define T^* as the set of pixels with unknown material.
for i in 1 : nrow(T^*) **do**
    for j in 1 : nrow($\chi_s^\star(t^{(i)})$) **do**
        Calculate standardised residual Std = $Std(\chi_{s(1,j)}^\star(t^{(i)}) - \mathbb{E}_{\chi_s(T')}[\eta(t^{(i)})])$.
        if Std > ImThresh **then**
            Remove element $\chi_{s(:,j)}^\star(t^{(i)})$ from $\chi_s^\star(t^{(i)})$.
        end if
    end for
end for
for i in nrow(T) **do**
    Set $\hat{x_S}(t^{(i)})$ as the sample mean of $\chi_s^\star(t^{(i)})$.
end for

but − in order to show that a result can be found with an extreme initial condition − the first proposal solution was chosen to be a cuboid entity with thickness 10 and alloy 0.2. Figure 5.15 shows the mean implausibility measure for every iteration and the mean standardised error. The implausibility decreases swiftly towards a non-implausible solution. The standardised error oscillates between over and under estimates. This is expected for this problem as an overestimation of scatter implies an underestimation of thickness and alloy fields which in turn implies an overestimation of scatter. This is not strictly true for complex samples but not a bad rule-of-thumb.

Algorithm 14 Iterative Scatter Removal in Simulation

Inputs: a, b, a_1, b_1, l, l_1, l_2, ImThresh, z

Outputs: $\hat{x_S}$, $\chi_s^\star$

1: Apply Algorithm 10 to create Prior-Specification-for-$x_S^\star$.

2: Using Prior-Specification-for-$x_S^\star$, design a look-up table and simulation design, notated X and $X^{\mathcal{LU}}$.

3: Using X, $X^{\mathcal{LU}}$, a, b, a_1, b_1, l, l_1 and l_2, apply Algorithm 6 to create the look-up-tables and Direct-Beam Emulator.

4: Using Prior-Specification-for-$x_S^\star$, $\mathcal{LU}$, Direct-Beam-Emulator, z and ImThresh, apply Algorithm 9 to find non-implausible point $\hat{x_S}$ and the sets $\chi_s^\star$.

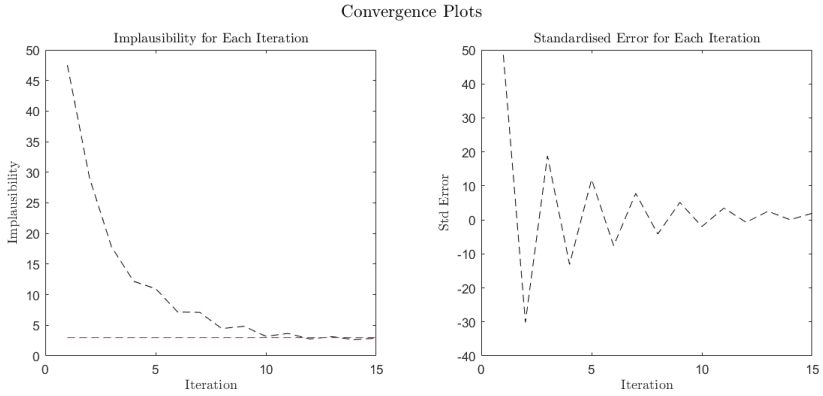

Figure 5.15: Left: implausibility measure for every iteration of the process. Right: standardised error for every iteration of the process.

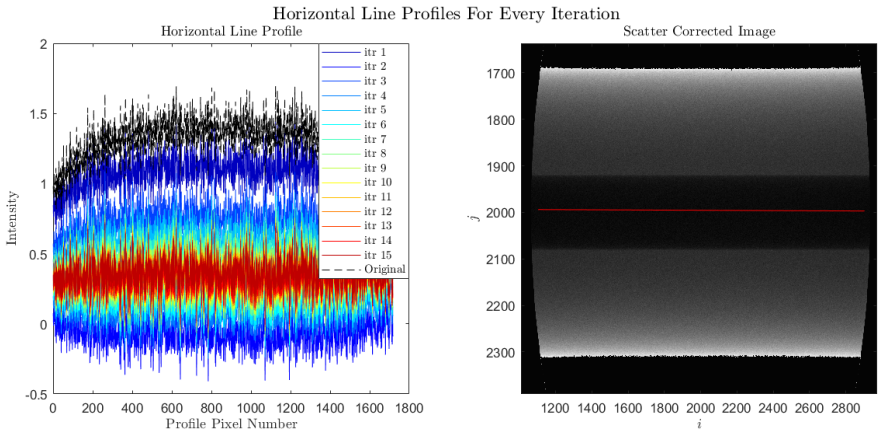

Figure 5.16: Left: intensity profiles for the scatter corrected image for every iteration. Right: simulated image with a red line depicting the profile showing the profile shown on the left.

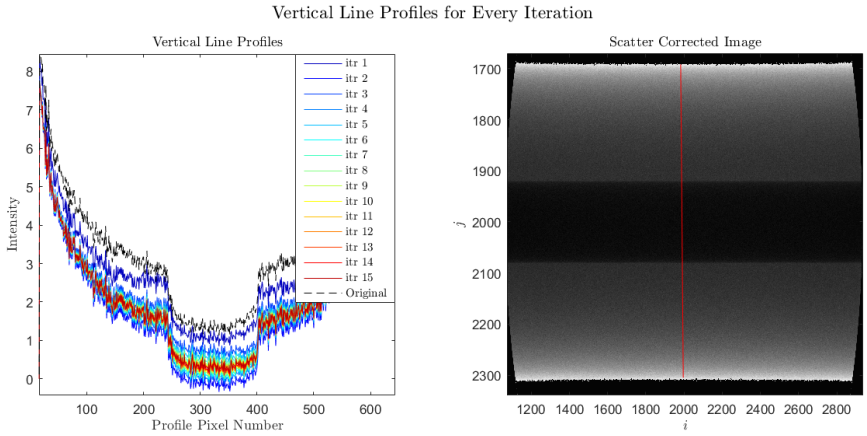

Figure 5.17: Left: intensity profiles for the scatter corrected image for every iteration. Right: simulated image with a red line depicting the profile shown on the left.

Figure 5.16 and 5.17 shows vertical and horizontal line profiles of the simulated image minus the expected scatter for every proposed solution. In order to return the required resolution, the expected scatter was re-scaled using a B-spline, see Appendix A.2. The solution swaps between over and under estimation for both horizontal and vertical profiles which agrees with the standardised error. The horizontal profile visibly flattens as the process moves closer to a non-implausible point which indicates that the process is moving towards an acceptable solution.

Figure 5.18 shows a set of summary outputs from the algorithm. The first image is the original image minus the expected scatter given the system value is the sample mean of the posterior non-implausible set. The two central images are the mean of the non-implausible sets of the thickness and alloy fields. This provides a diagnostic measure of effective thickness of aluminium which can be easily transformed to provide aBMD.

Figure 5.19 validates a non-implausible solution against known system values. Apart from some large errors near the edges, the percentage error is mostly in the region $[\pm 10]$. This validates that the expected entity composition can be used as a point estimate for material composition and the error is small enough to be useful to clinicians. The standardised errors are calculated using the non-implausible set given a non-implausible scatter estimate. The standardised errors are mostly between $[\pm 3]$ indicating that the uncertainty has been well estimated and there is no evidence that the known truth is not inside the non-implausible sets. We also checked that the known truth was within the non-implausible sets for the majority of sets in $\chi_s^\star$. This figure shows that $\chi_s^\star$ gives reasonable uncertainty estimation and the non-implausible point has a *small enough* percentage error.

Another output from this algorithm is a scatter corrected image. The scatter removal method is, given the system value is a non-implausible point, calculate i) $\mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}'(x_S^\star, M = 2)]$ ii) up-sample it using a B-spline and iii) take it away from the observation. Figure 5.20 is a plot illustrating the improvements in image quality when the expected scatter is removed. The images were normalised separately so that the image histogram peaks for the regions of interest overlap. In the left hand images, areas are saturated and there is large amount of vignetting, (see Section 4.3.1). This means that the fine detail shown by the fact that the local noise on the images is no longer visible. This is particularly obvious in the bottom image which is focusing on the region containing aluminium.

5.7 Conclusions

In this chapter, a method for making inferences for system values given an observation of system behaviour has been presented. The method finds non-implausible sets which may be used directly for decision making or to aid the calculation of adjusted beliefs. Even if the system values have a large number of dimensions, the method is able to efficiently find non-implausible sets using an exhaustive method or a non-implausible point using an iterative technique.

The exhaustive method is stable and robust but it is slow. The results presented took around 24 hours to run and that was for an entity with a simple geometry. As the entity complexity increases, so will the number of parameters needed to describe its material and thickness fields, making the process slower still. The iterative technique is fast but there is no guarantee that it will find a non-implausible point if one exists. Furthermore, there is less information about the uncertainty in the parameters. As the aim of this study was to provide a proof-of-concept for an algorithm, the goal of which is to infer entity composition and improve image quality without impeding clinical practice, the iterative method was the most feasible solution that could scale up to real world samples in the time frame of the study. The iterative technique has been tested on many entities in both simulation and the real world and finds a non-implausible solution if one exists. A subset of the simulated tests has been presented in this section and a subset of the real world tests will be presented in Section 6.5. So, although we are cautious about advocating its use in general, for the problems presented here, we believe it is effective and robust.

In Chapter 4, the advantages of our approach to predicting scatter – in that the error to the simulator was smaller – were presented. The advantages of our approach to making

inferences about scatter in an X-ray image from that image are: i) estimates for material and thickness (not just thickness) are returned and may be used to calculate diagnostic quantities and ii) we use an absolute (not relative) measure of whether an acceptable match to an X-ray image has been attained. Therefore, the advantages of our approach so far are: i) the method can predict scatter more accurately than current state-of-the-art approaches, ii) whether the solution is *good enough* is well tested and iii) diagnostic measures are returned from a mono-spectral X-ray image.

There are two principle directions for future research which we envisage may improve this method. The first is to use the iterative method to find an interesting area and then perform the exhaustive approach around that area. This will give a better representation of the non-implausible set. Furthermore, additional simulations may be run in the surrounding area reducing uncertainty in the simulator. The second direction is to perform a multi-resolution analysis. The speed of the analysis scales with the resolution of the detector squared. As the analysis at full resolution is impracticable, it is performed at a lower resolution and then the expected scatter is up-sampled. However, given that this is useful, it is logical that the lower resolution non-implausible sets are informative for a higher resolution version. Therefore, an exhaustive history match could be performed at low resolution. The low resolution non-implausible set is used to design a refocused non-implausible set at a slightly higher resolution. This is then refocused using an emulator built at that resolution. This could be repeated until the non-implausible set considered is at the resolution required to model the real world effectively. As the space in which the emulator needs to be accurate is smaller, it is likely this approach will also improve emulator efficiency. This could also be combined with a sequential history matching approach which might improve efficiency. A development which would enable these improvements is a more sophisticated spatial model which encodes an expert's beliefs about complex human anatomy.

The application has shown a proof-of-concept in simulation that, given G4 is the real world, even if the material is uncertain, i) scatter can be effectively removed, ii) image quality improved and iii) inferences made about the composition of the entity. However, conditioning an inference on the model being correct is not enough to be able to give these benefits to clinicians. The aim is to make inferences for real entities and real physical systems. This leads us onto the next chapter which focuses on making inferences for physical systems using complex computer models.

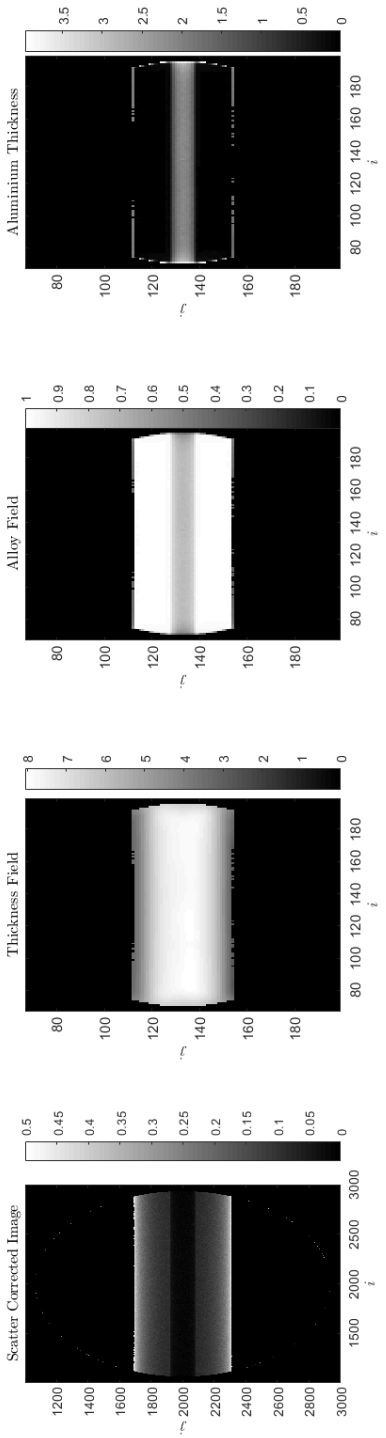

Figure 5.18: Top: original image minus expected scatter given a non-implausible point. Top centre: non-implausible alloy field. Bottom centre: non-implausible thickness field. Bottom: non-implausible thickness of aluminium.

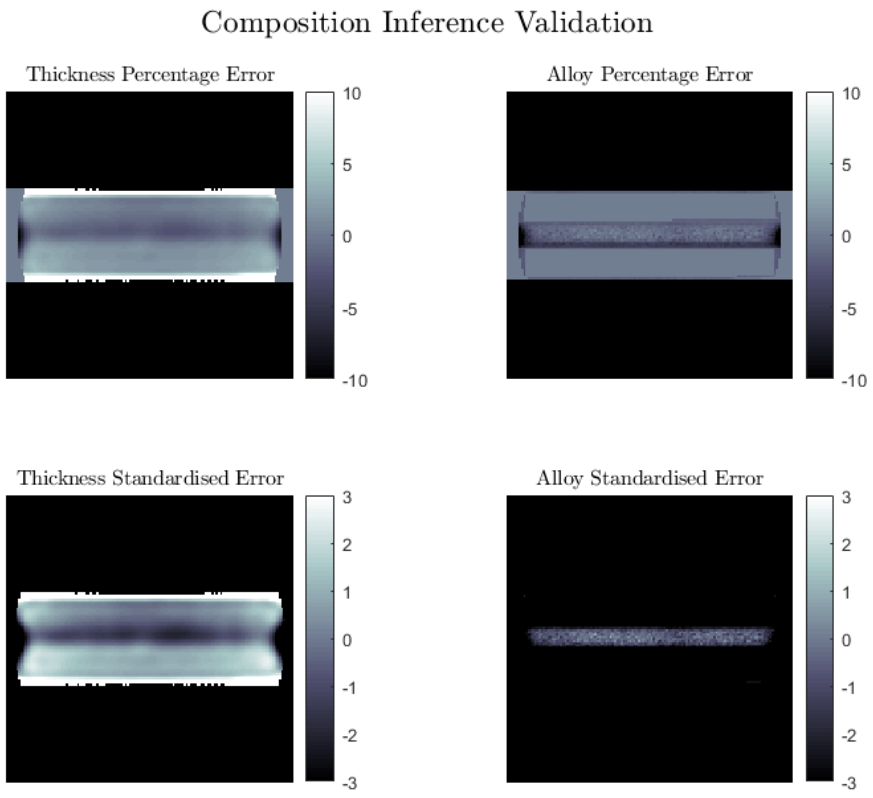

Figure 5.19: Top: Percentage error between non-implausible system values and known truth. Bottom: standardised errors between adjusted beliefs and known truth.

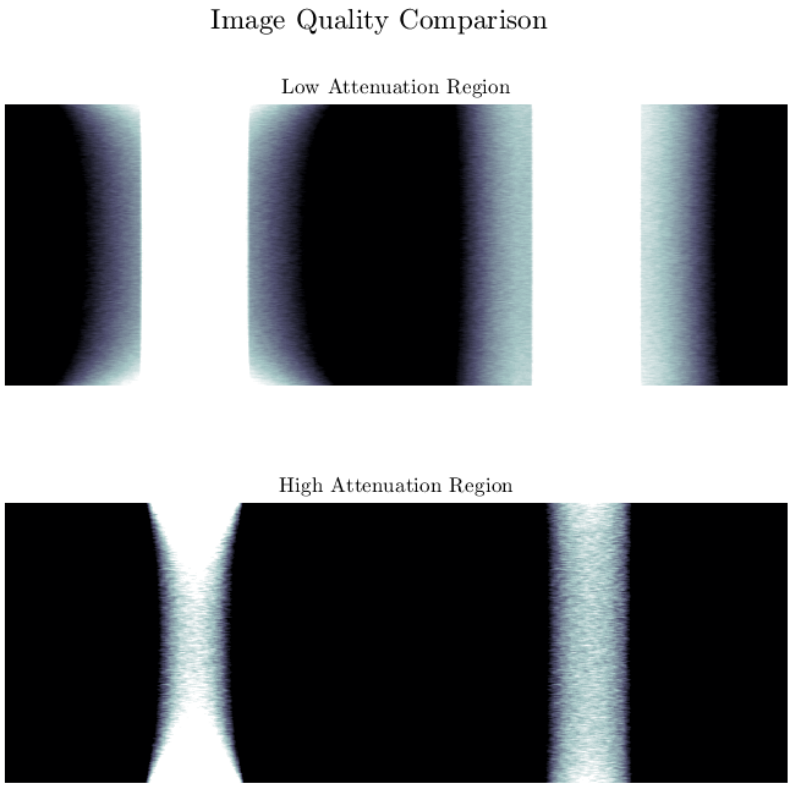

Figure 5.20: Top: image comparison of original (left) and scatter corrected (right) images with contrast stretched to compare the PMMA region. Bottom: image comparison of original (left) and scatter corrected (right) images with contrast stretched to compare the aluminium region.

Chapter 6

Reified Bayesian Emulation for Multiple Physical Systems

In this thesis, methods for making inferences for the scatter present in a simulated X-ray image, given system values, and making inferences about system values, given a simulated X-ray image, have been presented. The goal is not however to make inferences from simulated X-ray images but to be able to make inferences about physical phenomena from real world X-ray images. The problem is that the data-sets required for inferring entity composition and expected scatter are impracticable to acquire experimentally. This is primarily for three reasons: i) it is difficult to produce pencil beams in the real world, ii) it is not feasible to create the entities required for training and iii) the number of observations needed are prohibitively expensive to attain. Therefore, a method for using a simulated database to make inferences about physical systems is necessary. This chapter will propose our solution to problem statement 1.1.4, the transferable database problem.

In this thesis, a transferable database is a simulated database which may be used to make inferences about real world physical phenomena. The process is not to transfer a database and then use that to make inferences in the real world, but to use a simulated database to make inferences in the real world directly. The process has three steps i) create a look-up-table (developed in Section 4.3.4) ii) use that look-up-table to model divergent beam X-ray images and iii) transfer those images. The look-up-table is used to make inferences in the real world, but that database is never explicitly transferred. Once the convolution (see equation 4.3.7) is performed and an image created, that image can be transferred to the real world images. The method could be used to create a transferred database of images but, in this work, it is not. Transferable databases fall under the general

class of problems that aim to make inferences about physical systems using physics models. Many of these systems are analogous to X-ray imaging in that a single simulator can be considered as informative for multiple systems. A method for making inferences about multiple physical systems using a physics model is therefore of interest and will be the focus of this chapter.

Learning about physical phenomena using complex computer models poses a number of questions: i) what is the physical meaning of the simulator inputs and outputs, ii) how can the simulator's imperfections be quantified, iii) how is one simulator informative for multiple physical systems, iv) how are observations on multiple physical systems informative for a simulator's imperfections and v) how may improvements to the simulator be assessed?

Many attempts at a mathematical solution to the scattering problem use physics models, see [62, 64, 66, 67, 69, 71, 77, 78]. It was discussed in Chapter 4 that assuming PSFs are piece-wise-stationary and isotropic induces discrepancy between emulator and simulator. Some papers attempt to quantify this error either in simulation or by comparison to real world X-ray images. For example, [62] assesses this error in simulation while [123, 124] assess the accuracy of scatter prediction against measurements with an ASG. Another error which is not directly considered is that, even if a stationary-isotropic PSF is appropriate (for example a cuboid of PMMA), the simulator does not equal reality. Although it is accepted that there is discrepancy between a simulator and reality, there is no serious attempt to properly consider i) what is causing the error, ii) how it may be reduced or iii) how model discrepancy should be included in the inference. These three points are addressed in this chapter. The sources of discrepancy between our simulator and the actual X-ray examination system will be discussed. A method is developed to calibrate GEANT4 reducing the difference between our simulator and reality. Finally, the unresolved portion of the discrepancy is quantified and included in the criteria for an acceptable match in our history match. These features are novel in scatter estimation and enable further improvements over current state-of-the-art approaches to the scattering problem.

This chapter will be organised as follows. Section 6.1 will discuss the simple case of making inferences about physical systems using physics models and its extension to multiple systems. Section 6.1.2 will discuss reified Bayesian analysis. Section 6.2 and 6.2.2 will discuss tuning to a reified simulator. Section 6.3 will discuss structural tuning. Finally, Section 6.4 will apply the methods presented to making inferences in the real

world using GEANT4, see Section 4.3.2.

6.1 Simulators and Multiple Physical Systems

As simulators can be used to learn about phenomena that are impractical to analyse directly, they are effective tools to aid understanding about complex physical systems. For example, a simulator can be used to constrain forecasts to only those which obey physical laws and it is reasonable to think that this will improve a forecast of system behaviour. The corresponding physical experiments cannot be performed for a number of reasons: i) the phenomena is too small or too large to observe directly ii) there are ethical or resource limitations in making the observations or iii) the phenomena has yet to occur. For example in climate change research, the process about which an expert would like to make inferences has not yet occurred and the outcome cannot be inferred from current data alone. However, any analysis using a physics model should include the fact that models are imperfect.

Simulators are built to be informative for reality and they are often studied to be able to make statements about physical systems. However, if an inference is conditioned on the model being reality, the analyses are really making a statement about the model [115]. How a simulator is informative for a physical system is often more complex. The aim is to adjust an expert's beliefs about the system values and system behaviour using a simulator and an observation. In order to do this there are additional uncertainties − not present when analysing the model − that must be considered. A model is not reality and there will be a difference between a simulator and the system it purports to represent. This difference is referred to as model discrepancy and there are two types of discrepancy which are considered in this work: model imperfection and parametric uncertainty.

1. **Model imperfection.** A simulator can be described as a function encoding the way system values change system behaviour. This relationship is imperfect for at least three reasons. Firstly, the system values are not always fully understood and, even if they are, they are often too complex to encode mathematically. Secondly, the mathematics encoding how system values change system behaviour are similarly incompletely understood or simplified. Finally, even if both of the above are perfect, there will be simplifications in the numerical solver.

2. **Parametric uncertainty.** There is additional uncertainty induced by the inputs.

This uncertainty is due to fact that the system values associated with these inputs are uncertain or the relationship between these inputs and system values is unknown. For example, inputs may be related to the experimental set up and make the simulator relevant to the piece of equipment the observations were made on. In G4 there are a number of inputs controlling the spectrum from which photons are sampled. However, given the spectrum produced by the source is uncertain, the inputs relating to the spectrum should also be.

The fact discrepancy exists does not invalidate using physics models to make inferences. However, to be able to make meaningful statements about the real world, model discrepancy should not be ignored.

This chapter is returning to the abstract discussion in Section 4.1.1. Equation 4.1.1 defines model discrepancy as the difference between i) a function of a set of simulations at carefully chosen input-output combinations and ii) reality. This construction was construct as it corresponded with our method for emulating X-ray images as a summation of components of pencil beam simulations. Equation 4.1.1 is stating that we do not assume that a simulator mimics reality perfectly. Although important to accept that discrepancy exists, without extra structure, this statement is not particularly profound as any number may be written as any other number plus an error. This statement is more meaningful when structure is added for example, it is assumed that model discrepancy has expectation one. In this case, explicitly considering model discrepancy is necessary as it enables coherent real world inference using simulators and not considering it could have a large impact on the inferences performed. In this chapter, structure will be added to model discrepancy enabling quantification of its magnitude and resolution of uncertainty about its value.

Traditional approaches to inference with complex computer models centre around what is termed calibration. They aim to find an input which is best in the sense of maximising a likelihood, usually minimising an error to an observation. Then prediction is performed using a single evaluation of the simulator at that *special* input. The problem is that it is hard to argue − given the imperfections in the simulator and the possibly ill-posed nature of the minimisation problem − whether a minimum exists, whether it can be found and why it is special. The discrepancy between model and system is not directly considered and the failure of a model to mimic a system is attributed to measurement error and a vague acceptance that the model does not equal reality [23].

There is a growing field in subjective Bayesian inference focusing on making inferences about physical systems using complex computer models. The pivotal development is that the methods explicitly include model discrepancy. Although model discrepancy is informally recognised, formally including it in the inference is a clear improvement on not including it. As judgements about the difference between a simulator and a system are often the judgements of experts, subjective Bayes is the appropriate paradigm for such an analysis. An early paper in which many of the uncertainties are described and a single simulator is used as informative for multiple systems is [31]. [32,34,36,41] apply subjective Bayesian methodologies to a range of physical systems using full Bayes, Bayes linear or a combination of the two. The approaches mostly follow the best-input assumption.

Definition 6.1.1. ***Best-input assumption.*** *There exits an input $x^\star$ where $x^\star \perp\!\!\!\perp \mathcal{F}$, such that $\mathcal{F}(x^\star)$ is sufficient for learning about the physical system [22].*

In other words, being sufficient for learning about the physical system means, if the best-input were known, no further evaluations of the simulator would change our beliefs about the system. Given the best-input assumption is asserted, the model discrepancy is independent of $\{\mathcal{F}, x^\star\}$. Given there is measurement error and model discrepancy, it is accepted that the best-input cannot be identified precisely and inference is performed using Bayesian inverse or prediction methodologies. This gives a method for uncertainty quantification in both best-input and prediction of system behaviour.

The two papers [23, 24] provide a coherent approach for making statements about physical systems using complex computer models. They carefully define i) inputs, outputs and system behaviour, ii) a coherent method for making inferences when inputs cannot be directly associated with physical quantities and iii) an approach when multiple simulators of the same physical system are available. The methods can involve a high level of elicitation which makes them difficult as elicitation methods are currently lagging behind their computational counterparts [103]. It is however clear the physical quantities about which they are making inferences and the logical basis for believing the inference they perform is applicable to those quantities.

Here a combination of i) reified Bayesian analysis [23], ii) the tuning approach of [24] and iii) inference for multiple physical systems of [31] will be developed. Particularly we will concentrate on how observations of the behaviour of multiple systems − for which the system values are well understood − may be used to make inferences about model discrepancy for similar physical systems. In this work, changes to $\mathcal{F}$ or the input space of

$\mathcal{F}$ will be considered as a new simulator. The set of simulators informative for the same physical systems is referred to as a family of simulators. For example, consider a simulator $\mathcal{F}(x)$, if a parameter w fixed in $\mathcal{F}$ is released, a new simulator in the family is defined as $\mathcal{F}'(x, w)$.

6.1.1 General Form: the Simple Case

The simplest case of simulator inference is when it is assumed that the output is the system behaviour and the inputs are measurable.

Definition 6.1.2. ***Measurable inputs:*** *inputs which correspond to system values which can, in principle, be measured independently of the simulator [24].*

Asserting the best-input is measurable leads to the direct simulator assumption.

Definition 6.1.3. ***Direct simulator assumption:*** *$x^\star$ is measurable and $\mathcal{F}(x^\star)$ is sufficient for the simulator for learning about y [24].*

If the relationship between simulator and system behaviour is given by equation 5.1.1, this assumption implies ϵ_{md} has expectation zero and is independent of $\{x^\star, \mathcal{F}\}$. The model discrepancy is defined as the difference between the simulator evaluated at the system values and the system behaviour $\epsilon_{md} = y - \mathcal{F}(x^\star)$. In this case, it is clear what model discrepancy is; $x^\star$ can be measured and then the simulator evaluated to reveal ϵ_{md}. Model discrepancy is considered and fits tractably into the analysis as an additional source of uncertainty adjusting the criteria for an acceptable match to observations and increasing the uncertainty in system behaviour. The practicalities may be difficult but it is clear what physical quantities beliefs are being adjusted for and the logical reason for doing so. For example, specifying beliefs about $x^\star$ is specifying beliefs about the system values.

There are implications of asserting the direct simulator assumption. Consider the example of material inference using G4. In this example, $x^\star \perp\!\!\!\perp \epsilon_{md}$ means that we expect our simulator to reflect reality equally well, no matter the entity complexity or the material present. Do we think the simulator will perform equally well when simulating a complex alloy like bone or a simple one like aluminium, probably not. A thought experiment to illuminate another problem with this assumption, similar to that proposed in [23], is as follows. Consider two simulators $\mathcal{F}'(x, w)$ and $\mathcal{F}(x)$ where fixed parameters w in $\mathcal{F}(x)$ have been released to create $\mathcal{F}'(x, w)$. An expert should prefer $\mathcal{F}'(x, w)$ over $\mathcal{F}(x)$ as, for

some value w_0, $\mathcal{F}'(x, w_0) = \mathcal{F}(x)$. Therefore, if the direct simulator assumption is asserted for $\mathcal{F}$, it should also be asserted for $\mathcal{F}'$. The relationship between the simulators and the system values is given by

$$\mathcal{F}'(x^\star, w^\star) \oplus \epsilon'_{md} = y \text{ and}$$
$$\mathcal{F}(x^\star) \oplus \epsilon_{md} = y$$

where ϵ'_{md} is the model discrepancy associated with $\mathcal{F}'$ evaluated at system values $(x^\star, w^\star)$. Given that the model discrepancy is partly due to fixing w to w_0, $\mathcal{F}'$ will be informative for ϵ_{md}. For example, a sensitivity analysis could be performed to adjust beliefs about $\mathbb{V}ar(\epsilon_{md})$. However, as $\mathcal{F}'(x, w_0) = \mathcal{F}(x)$, $\mathcal{F}$ is informative for $\mathcal{F}'$. This implies $\mathcal{F}$ is also informative for ϵ_{md} contradicting the statement that $\mathcal{F}$ is a direct simulator. In many cases, even if the inputs are measurable, the model discrepancy is dependent on simulator and system values in a complex way. Furthermore, inputs often do not have a clear physical analogue. This may be because they are a hypothetical mathematical construct or because the resolution of the input is insufficient to describe the system values meaningfully. Without this analogue there are philosophical complications: why should an expert be able to specify beliefs about a quantity without a physical meaning and why would anyone have confidence in prediction if $x^\star$ is not the system values?

Multiple Systems: the Simple Case

Often the interest is in making inference about a single system with system values and behaviour, of which multiple observations have been made. In this situation, the simulator is generally designed with a specific context in mind. For example, a climate change simulator is designed to model Earth's climate and will have hard coded components specific to Earth. Other computer models are designed to be used in a more general context where changing the parameter combination makes the simulator informative for a specific context. A more general simulator of climate would be valid for many planets while a specific input combination makes it informative for Earth. There is no reason why a climate simulator would be made more generic if the interest is only in Earth's climate. In other situations, it is useful to design generic simulators. For example, $x^\star$ is the chemical composition of a particular drug and the simulator may model how many drugs are processed by the body. This situation poses interesting questions about how a single simulator may be informative for multiple physical systems and how multiple

physical systems can be informative for a simulator's imperfections.

Here two physical systems are considered distinct if the system values or behaviour need not be the same. In this work, the system is defined by the entity and the experimental setup $\{f, S, x_E\}$, see equation 2.2.3. Any time the experimental equipment, experimental conditions or entity examined is changed, it constitutes a new physical system. The simplest case, of the approach used in this thesis to model multiple physical systems, is to set up a system of linked equations. The direct simulator assumption is made for all relationships i.e. the system takes the form

$$\begin{aligned} \mathcal{F}(x^{\star(1)}) \oplus \epsilon_{md}^{(1)} \oplus \epsilon_{me}^{(1)} &= z^{(1)}, \\ &\vdots \\ \mathcal{F}(x^{\star(N)}) \oplus \epsilon_{md}^{(N)} \oplus \epsilon_{me}^{(N)} &= z^{(N)}. \end{aligned}$$

For an observation of a new system $z^{(k)}$ such that $k > N$, with associated system value $x^{\star(k)}$, how the system values relate to one another must be specified. In a Bayes linear analysis, a second order specification must be made for the set $\{x^{\star(k)}, x^{\star(1)}, \cdots, x^{\star(N)}\}$. For example, the set could be considered second order exchangeable (see Section 3.2). Beliefs must also be specified for the set of model discrepancies. For example, they may be specified as uncorrelated if their population mean is known, or part of an infinite second order exchangeable sequence if it is not. The difficulty comes when the direct simulator assumption is not commensurate with an expert's beliefs and, given the strict requirements of the assumption, it is likely that it is not.

There are features of this case which are worth further discussion. A natural assumption on discrepancy would be that the covariance between discrepancies would depend on the proximity of their associated system values. For example, the correlation could depend on $|x^{\star(i)} - x^{\star(j)}|$. In the most extreme case using two observations, we would expect that we would learn more about $\epsilon_{md}^{(i)}$ when $x^{\star(i)} = x^{\star(j)}$ than when $x^{\star(i)} \neq x^{\star(j)}$. However, the direct simulator assumption implies that we cannot assert even such a simple structure on the discrepancy which illustrates how strong an assumption it is.

6.1.2 General Form: Reified Bayesian Analysis

Reified Bayesian analysis is a method for making inferences about real world phenomena using simulators when the direct simulator assumption is not commensurate with an expert's beliefs. To motivate the need for reified Bayesian analysis, consider the following

simple example. An expert realises that the units in simulation and the observation are different. Given this realisation, it is easy to fix the problem by multiplying the simulator output by a scalar and it seems inappropriate to simply calibrate a simulator by varying the inputs. Reified Bayesian emulation is a method for dealing with analogous, but potentially more complex, inconsistencies arising from careful consideration of simulators. Although how we deal with such inconsistencies may be more complex, the idea that we should ignore them and calibrate by varying the inputs seems as inappropriate as if the discrepancies were simply caused by a difference in units.

The reified approach proposes that the expert considers a hypothetical simulator which has been improved to such an extent that they are happy to assert that it is direct. In other words, a simulator which fulfils its purported function, that is, represents an expert's beliefs about the system behaviour or exhausts an expert's beliefs about how their simulator may be improved.

Definition 6.1.4. ***Reified simulator.*** *A simulator $\mathcal{F}^\star$ is reified when it is asserted that*

$$y = \mathcal{F}^\star(x^\star, w^\star) \oplus \epsilon^\star_{md} \tag{6.1.1}$$

where $x^\star$ and $w^\star$ are system values and $\epsilon^\star_{md}$ is independent of $\{\mathcal{F}, \mathcal{F}^\star, x^\star, w^\star\}$ [23].

The simulator $\mathcal{F}^\star$ is not considered the *best* possible simulator but one which is good enough that the direct simulator assumption is coherent. The reified simulator is a hypothetical construct, but it is still possible to construct a statistical model for it. Given that the simulator available and the reified version explain the same phenomena, it is reasonable to expect that the simulator available is informative for its reified counterpart. The simulator $\mathcal{F}$ is used to make inferences for an emulator of $\mathcal{F}^\star$ and that emulator is used to make inferences for the physical system via equation 6.1.1.

We specify how the actual simulator $\mathcal{F}$ is informative for its reified counterpart $\mathcal{F}^\star$ with emulators. Consider that the emulator for the actual simulator and the reified simulator are specified as multivariate random fields of the form

$$\mathcal{F}(x) = \mathcal{M}(x) \oplus \mathcal{W}(x) \text{ and}$$
$$\mathcal{F}^\star(x, w) = \mathcal{M}^\star(x, w) \oplus \mathcal{W}^\star(x, w),$$

then a probabilistic specification for the sets

$$\{\mathcal{M}^\star, \mathcal{M}\} \text{ and } \{\mathcal{W}^\star, \mathcal{W}\} \tag{6.1.2}$$

encodes the relationship between them.

To reify is to make the abstract more real and to reify a simulator is to make it more real [1]. The idea behind the method is a powerful one that is prominent in subjective Bayesian inference; consider what an expert actually wants to learn about and what the data available tells them about that. To enable tractable inference, a direct simulator assumption is required and, if a thought experiment reveals a sure improvement to the actual simulator, it is incoherent to assert that it is direct. To be able to perform inference, an expert considers a simulator that is direct and what the actual simulator tells them about the direct simulator. To obtain this coherence, an expert must specify the relationship between their simulator and their reified counterpart. Eliciting these beliefs is difficult as the things about which an expert would like to make inferences are complex. However, reified analysis offers a route to coherently linking abstract simulators and the physical world.

6.1.3 Structural Reification

Structural reification is a method for specifying beliefs about the sets in equation 6.1.2. It asks an expert to considers how possible improvements to $\mathcal{F}$ might change its emulator. There are two prominent types of improvements: i) fixed inputs are made uncertain and ii) additional inputs are added to explain physical processes simplified or left out. For every improvement, an improved simulator is considered in which the previous simulator is nested. Then, an improved simulator may be written $\mathcal{F}(x) = \mathcal{F}'(x, w_0)$ and the relationship between $\mathcal{F}$ and $\mathcal{F}'$ may be written as

$$\begin{aligned}\mathcal{F}'(x,w) &= \mathcal{F}(x) + \left(\mathcal{M}^{+}(x,w) \oplus \mathcal{W}^{+}(x,w)\right) \\ &= \mathcal{M}'(x,w) \oplus \mathcal{W}'(x,w)\end{aligned}$$

where $\mathcal{M}^{+}(x, w_0) = \mathcal{W}^{+}(x, w_0) = 0$. When $\mathcal{F}$ is nested in $\mathcal{F}'$, beliefs about $\mathcal{F}^{\star}$ are separated from $\mathcal{F}$ by $\mathcal{F}'$, meaning, an expert's beliefs about the reified emulator are independent of the actual simulator given the improved simulator. Building the improvements in this way gives a structure to how many improvements may affect the actual simulator and thus beliefs about the reified simulator elicited.

[1] The term reify is often used as a negative. For example, "You have JUST reified your model.". In this context, reified is meant as you have taken an abstract concept to be real.

In this thesis, only nested simulators are considered. There is no reason in general that improvements should or could be nested in this way. However, if the improvements can be constructed using this form, the analysis is much easier as improvements may be considered sequentially making the elicitation easier and the task as a whole less daunting.

6.1.4 Reified Analysis for Multiple Systems

The novel extension of reified analysis to multiple systems follows on from the simple case. The observations are related to the reified simulator via the direct simulator assumption for every system

$$\begin{aligned}\mathcal{F}^\star(x^{\star(1)}, w^{\star(1)}) \oplus \epsilon_{md}^{\star(1)} \oplus \epsilon_{me}^{(1)} &= z^{(1)},\\ &\vdots\\ \mathcal{F}^\star(x^{\star(N)}, w^{\star(N)}) \oplus \epsilon_{md}^{\star(N)} \oplus \epsilon_{me}^{(N)} &= z^{(N)}.\end{aligned}$$

Here the set $\{\epsilon_{md}^{\star(1)}, \cdots, \epsilon_{md}^{\star(N)}\}$ is uncorrelated and the reified simulator is related to the actual simulator via

$$\begin{aligned}\mathcal{F}^\star(x^{\star(1)}, w^{\star(1)}) &= \mathcal{F}(x^{\star(1)}) \oplus \Delta(x^{\star(1)}, w^{\star(1)}),\\ &\vdots\\ \mathcal{F}^\star(x^{\star(N)}, w^{\star(N)}) &= \mathcal{F}(x^{\star(N)}) \oplus \Delta(x^{\star(N)}, w^{\star(N)}).\end{aligned}$$

In this thesis, Δ is referred to as a transfer function to the reified case. The terminology was developed as a compromise between experts wanting to transfer the simulator to reality and the notion that a reified simulator is the realistic best case. For a further physical system with system values $\{x^{\star(k)}, w^{\star(k)}\}$ where $k > N$, the additional specifications required are for the set

$$\{x^{\star(k)}, w^{\star(k)}, x^{\star(1)}, w^{\star(1)} \cdots, x^{\star(N)}, w^{\star(N)}\}$$

and the transfer function $\Delta(x, w)$. When considering a single system, the observation is only considered informative for $\mathcal{F}^\star(x^\star, w^\star)$ and not $\mathcal{F}^\star$ in general. This is because the relationship is complex and it is likely to have a small impact on the inferences performed. When considering multiple systems, there are many evaluations of the reified simulator which are sufficient to describe many physical systems' behaviours. Suppose a set of observations could be made for which the system values are effectively known. For example, a set of X-ray images of objects which have been manufactured with a specified material and a specified geometry. Then the difference between the observations and the actual

simulator is given by

$$z^{(i)} - \mathcal{F}(x^{\star(i)}) = \Delta(x^{\star(i)}, w^{\star(i)}) \oplus \epsilon_{md}^{\star(i)} \oplus \epsilon_{me}^{(i)}$$

which is informative for the difference between the actual simulator and its reified counterpart. Therefore, when multiple systems are considered, observations can be used to resolve an important portion of the uncertainty in the reified simulator by resolving uncertainty in $\Delta(x, w)$.

6.2 Tuning to a Reified Simulator

The novel method used in this thesis for specifying the relationship between the actual simulator and its reified counterpart is tuning. The inputs are split into $x = (x_S, x_E)$. The subscripts denote sample (entity) and equipment to coincide with the X-ray application but they could be referred to as measurable and tuning as in [24]. x_S are the inputs relating to system values of direct interest and x_E are those only of interest to the extent that knowledge of their value will reduce uncertainty in $x_S^\star$ and y. Referred to as tuning inputs, they are allowed to drift from their measurable counterparts in order to resolve some of the model discrepancy. The relationship between a simulator and its reified version is given by

$$\mathcal{F}(x_S, x_E') + \Delta(x_S, x_E') = \mathcal{F}^\star(x_S, x_E^\star, w^\star). \tag{6.2.1}$$

The approach proceeds by making a best-input assumption on x_E' (see definition 6.1.1) in that it is assumed that there exists a single value of x_E such that, if it were known, it is sufficient for the actual simulator for learning about the reified simulator. Tuning is useful when the system properties in the actual simulator are unidentifiable from those in $w^\star$ and the difference between $\mathcal{F}$ and $\mathcal{F}^\star$. For example, in X-ray imaging, there is likely to be a number of materials which make up the filtration at the source. However, all that is of interest is the spectrum of photons that is incident on the entity. Given that the same spectrum can be produced using one thickness of aluminium, it seems reasonable to allow the thickness of aluminium at the source to drift to account for all the other materials present. This pragmatic decision makes the analysis more tractable by reducing the number of inputs.

It is not believed that the tuning input actually exists. It is a hypothetical construct which enables tractable analysis. Therefore, if it is to be used in an inferential framework, whether it is an acceptable modelling assumption should be assessed. History matching

provides a route to such an analysis returning the set of tuning inputs for which the error induced by the assumption is small compared to the current level of uncertainty. The problem is that if it is a hypothetical construct, it may be difficult for an expert to specify the beliefs required for a history match. For example, specifying plausible bounds for the tuning parameter.

6.2.1 Tuning for Multiple Systems

To apply tuning to multiple systems, it is assumed that a set of calibration observations, notated $D^f = \{z^{(1)}, \cdots, z^{(N)}\}$, is available and the associated entity dependent system values are known. Assuming the entity dependent system values are known for the calibration entities is reasonable for our example as entities can be constructed with known dimensions and chemical compositions. The system of calibration measurements are linked by

$$\begin{aligned} \mathcal{F}(x_S^{\star(1)}, x_E^{\prime(1)}) + \Delta(x_S^{\star(1)}, x_E^{\prime(1)}) \oplus \epsilon_{md}^{\star(1)} \oplus \epsilon_{me}^{(1)} &= z^{(1)}, \\ &\vdots \\ \mathcal{F}(x_S^{\star(N)}, x_E^{\prime(N)}) + \Delta(x_S^{\star(N)}, x_E^{\prime(N)}) \oplus \epsilon_{md}^{\star(N)} \oplus \epsilon_{me}^{(N)} &= z^{(N)}. \end{aligned}$$

Here $z^{(i)} = f^{(i)}(S^{(i)}, x_E^{f(i)}) \oplus \epsilon_{me}^{(i)}$ (see equation 2.2.3) and the relationship between the reified case and the systems behaviours is specified by

$$\begin{aligned} \mathcal{F}^\star(x^{\star(1)}, w^{\star(1)}) \oplus \epsilon_{md}^{\star(1)} &= f^{(1)}(S^{(1)}, x_E^{f(1)}), \\ &\vdots \\ \mathcal{F}^\star(x^{\star(N)}, w^{\star(N)}) \oplus \epsilon_{md}^{\star(N)} &= f^{(N)}(S^{(N)}, x_E^{f(N)}). \end{aligned}$$

It is also assumed that $x_S^{\star(i)} \perp\!\!\!\perp \{f^{(j)}, x_E^{f(j)}\}$, $x_S^{\star(i)} \perp\!\!\!\perp x_E^{\star(j)}$ and $x_E^{\star(i)} \perp\!\!\!\perp S^{(j)}$ $\forall i, j$. This assumption is that the entity itself is independent of the experimental system used to examine it. That is not to say the level to which uncertainties may be resolved are are not affected by the equipment, only that knowing the equipment *a-priori* gives no information about the entity.

Direct Tuning Assumption (Entity Independent Transfer Function)

There are a wide range of beliefs which may be encoded in this relationship. For example, the set of tuning inputs could be specified as second order exchangeable. However, in this thesis it is assumed i) that the difference between tuned simulator and reified simulator is

only dependent on the tuned simulator output and ii) that the equipment and experimental settings are constant. This is referred to as the direct tuning assumption, given by

$$\begin{aligned} &\Delta \perp\!\!\!\perp \{\mathcal{F}, x_S^\star, x'_E, w^\star, x_E^\star, \mathcal{F}^\star\} | \{\mathcal{F}(x_S^\star, x'_E)\}, \\ &x_S^{\star(k)} \perp\!\!\!\perp \{x_S^{\star(1)}, \cdots, x_S^{\star(N)}\} \; \forall k > N \text{ and} \\ &x_E^{\prime(i)} = x_E^{\prime(j)} \; \forall \; i, j. \end{aligned} \tag{6.2.2}$$

The assumption that the transfer function is independent of the system values and simulator, given the simulator evaluated at the tuning input, is pivotal to a tractable solution to problem statement 1.1.4. The assumption is that if

$$\mathcal{F}^\star_{(t)}(x_S^{\star(i)}, x_E^\star, w^\star) = \mathcal{F}^\star_{(t)}(x_S^{\star(j)}, x_E^\star, w^\star),$$

then there is an input x'_E such that

$$\mathcal{F}_{(t)}(x_S^{\star(i)}, x'_E) = \mathcal{F}_{(t)}(x_S^{\star(j)}, x'_E) + \epsilon \tag{6.2.3}$$

for all $x_S^{\star(l)}$, where ϵ is small enough − when compared to other sources of variation − that it may be ignored without a large impact on the inferences performed. This assumption enables tractable prediction for measurable inputs outside the convex hull of the measurable inputs within the calibration training-set. When considering our X-ray application, we would like to make inferences using a simulated database and have available a set of observations of cheap, malleable and readily available materials. For example, we could use cuboids of PMMA of which the dimensions are known. If it can be assumed that the difference between simulator and reified counterpart is independent of the material given the tuned simulator output, the cheap malleable materials can be used to learn about the transfer function and tuning parameters. Then, inferences may be made for the reified simulator for any material and direct simulator inference performed. We term this a material invariant transfer function in our X-ray image example or an entity independent transfer function more generally. Ensuring that it is reasonable to assert the transfer function is material invariant is a key consideration in this work.

The fact that this assumption is acceptable is not obvious or general. Indeed, we were sceptical that it could be asserted at all in our X-ray example. However, as will be presented in the following sections, analysis can be constructed so that whether it is acceptable is carefully validated. If this cannot be asserted, analysis can be performed to find a simulator for which it can. If no such simulator exists, a general uncertainty

analysis, possibly using the methodology set out in [23], would have to be performed. To show that − at least in one case − it is acceptable, consider the simple simulator

$$\mathcal{F}_{(t)}(x_S, x_E) = x_S x_E t \text{ and } \mathcal{F}^{\star}_{(t)}(x_S, x_E, \lambda) = \lambda \mathcal{F}_{(t)}(x_S, x_E).$$

This specification of the relationship between actual and reified model might be appropriate if the reified version and the actual simulator have different scales. As any value of λx_E can be attained by allowing x_E to become a tuning parameter, if λ and x_E are not of direct interest, it may not be worthwhile adding λ explicitly to the analysis. Furthermore, for this example, the predictive properties are assured even when tuning. The cost of making this assumption is that the meaning of λ and x_E is lost and a more careful argument that the tuned simulator has good predictive properties is required.

6.2.2 Inference Using the Direct Tuning Assumption

Given that the direct tuning assumptions have been asserted, the next step is to make inferences from the available data. The random constructs of interest are the tuning parameters and the transfer function. The aim is to history match both tuning parameters and transfer functions to find a sample from the set which provides an acceptable match to observed data. In other words, the set of tuning parameters and transfer functions which do not contradict the assumptions expressed in relation 6.2.2 and give an acceptable match to observed data. The idea is similar to conventional history matching: i) define a prior non-implausible set for x'_E, ii) for each value in the set adjust beliefs about the transfer function given $x_E = x'_E$ and iii) calculate an implausibility measure.

The set of calibration measurements may be rewritten in the form

$$\begin{gathered}\Delta(\mathcal{F}(x_S^{\star(1)}, x'_E)) \oplus \epsilon_{md}^{\star(1)} \oplus \epsilon_{me}^{(1)} = z^{(1)} - \mathcal{F}(x_S^{\star(1)}, x'_E) = \epsilon^{\Delta(1)}, \\ \vdots \\ \Delta(\mathcal{F}(x_S^{\star(N)}, x'_E)) \oplus \epsilon_{md}^{\star(N)} \oplus \epsilon_{me}^{(N)} = z^{(N)} - \mathcal{F}(x_S^{\star(N)}, x'_E) = \epsilon^{\Delta(N)}\end{gathered}$$

and used to adjust beliefs about the transfer function and the tuning parameter x'_E. Beliefs about the transfer function are specified as a multivariate random field in the form

$$\Delta_{(t)}(\mathcal{F}(x_S^{\star}, x'_E)) = \sum_m \beta_{(m,t)} h_{(m)}(\mathcal{F}_{(t)}(x_S^{\star}, x'_E)) \oplus \mathcal{W}^{\Delta}_{(t)}(\mathcal{F}_{(t)}(x_S^{\star}, x'_E)).$$

Then an expert makes a second order specification for β and $\mathcal{W}^{\Delta}$. If an expert is unable to make the specifications, GLS estimates for adjusted beliefs (see key formula 2.2.1) may

be used instead. The quantities $\epsilon^{\Delta(i)}$ are informative for the transfer function and the second order specification is given by

$$\begin{aligned}
\mathbb{E}[\epsilon^{\Delta(i)}] &= \mathbb{E}[\Delta(\mathcal{F}(x_S^{\star(i)}, x'_E))], \\
\mathbb{V}ar[\epsilon^{\Delta(i)}] &= \mathbb{V}ar[\Delta(\mathcal{F}(x_S^{\star(i)}, x'_E))] + \mathbb{V}ar[\epsilon_{me}^{(i)}] + \mathbb{V}ar[\epsilon_{md}^{\star(i)}], \\
\mathbb{C}ov[\epsilon^{\Delta(i)}, \epsilon^{\Delta(j)}] &= \mathbb{C}ov[\Delta(\mathcal{F}(x_S^{\star(i)}, x'_E)), \Delta(\mathcal{F}(x_S^{\star(j)}, x'_E))] + \mathbb{C}ov[\epsilon_{me}^{(i)}, \epsilon_{me}^{(j)}] \text{ and} \\
\mathbb{C}ov[\epsilon^{\Delta(i)}, \Delta(\mathcal{F}(x_S^{\star(j)}, x'_E))] &= \mathbb{C}ov[\Delta(\mathcal{F}(x_S^{\star(i)}, x'_E)), \Delta(\mathcal{F}(x_S^{\star(j)}, x'_E))].
\end{aligned}$$

With this approach, there is a danger of over-fitting. This means that the transfer function resolves too much variation, all implausibilities are all too small and whether the input is consistent with the assumptions expressed in equation 6.2.2 is not assessed. If 6.2.2 does not hold, prediction outside the convex hull of the training-set is far from assured.

To ensure that the modelling assumptions are being tested by the history matching approach, instead of adjusting beliefs about the transfer function using the entire calibration training-set, it is split into two subsets $D^f = \{D_1^f, D_2^f\}$. The convex hulls of these two subsets are designed so that they are mutually exclusive and the observations in both groups span a similar range. For example, in our X-ray application, both sets could be distinct materials with thicknesses chosen so that the simulator output spans the same range. Then a given point x_E is classified as implausible if there is evidence to suggest that there is discrepancy between the two sets of adjusted beliefs or the observations are not acceptably well represented. Assuming the simulator has a scalar output, the implausibility measure used for this purpose is

$$\max_i \left\{ \left| \frac{z^{(i)} - 0.5(\mathbb{E}_{D_1^f}[\mathcal{F}^\star(x_S^{\star(i)}, x_E^\star, w^\star); x'_E = x_E] + \mathbb{E}_{D_2^f}[\mathcal{F}^\star(x_S^{\star(i)}, x_E^\star, w^\star); x'_E = x_E])}{\sqrt{\mathbb{V}ar[\epsilon_{me}^{(i)} + \epsilon_{md}^{\star(i)}] + 0.25(\mathbb{V}ar_{D_1^f}[\Delta(x_S^{\star(i)}, x'_E); x'_E = x_E] + \mathbb{V}ar_{D_2^f}[\Delta(x_S^{\star(i)}, x'_E); x'_E = x_E])}} \right| \right\}. \tag{6.2.4}$$

Here the numerator is the observation minus the average of the adjusted expectation for the reified emulator where each adjustment is performed with a distinct training-set. This is small when the transfer functions are similar and both mimic observed data well. Therefore, this jointly tests assumption 6.2.2 and whether an acceptable match to observed data has been attained. The denominator has the uncertainty due to measurement error, model discrepancy and the fact that the transfer function is not known but has been inferred from the data. There is a covariance term missed from the denominator but, in our X-ray example, testing showed that this did not have a large impact of the inferences performed.

Diagnostics

As with any inferential procedure, diagnostics are required to validate the inference. Here an expert's current uncertainty and the tuning assumption itself must be assessed. This may be performed using validation entities or cross validation if resources are available. For a validation entity − or an entity removed from the belief adjustment for cross validation purposes − the process is i) find the implausibility of the validation observation and ii) check it is less than a specified threshold. For example, if Pukelsheim's three sigma rule is used [47], check that the implausibility is less than three.

There are a number of features which indicate that the specifications require further consideration. The first is when the distribution of implausibilities has too many small or high values. This indicates poor uncertainty estimation or the assumption of independence on the transfer function being inappropriate. The second is correlation between implausibility and known system values. This again indicates a failure of the independence assumption of the transfer function. If many values are above the implausibility threshold or these features are obviously present, there is evidence to suggest that reconsidering the direct tuning assumption may be beneficial.

6.3 Structural Tuning

Reified Bayesian analysis was constructed as a coherent method for making inferences about physical systems with complex computer models. Allowing parameters to vary from their system values is a powerful tool for resolving discrepancy. However, the coherence of reified approach is lost as i) the tuning inputs are not operationally defined so it may be unreasonable for an expert to specify beliefs about them, ii) in general a tuned simulator need not give rise to an acceptable match to a reified simulator and iii) it is often hard to argue why the tuned simulator has good predictive properties if it is not evaluated at the system values. In this section, a novel method is presented which can be used to i) increase confidence that inferences performed using the tuned simulator are relevant in the real world and ii) for specifying beliefs about the random quantities involved in the analysis.

The aim of this section is to set up a methodology for i) assessing a-priori whether we expect our simulator to be tunable, ii) specifying beliefs about the quantities involved which are not operationally defined and iii) making second order specifications for diag-

nostic quantities which may be calculated when real world observations become available. The approach was designed to be analogous to a Bayes linear analysis where i) a second order specification is made and the expected adjustments analysed, ii) observations are made and the actual adjustments calculated and iii) comparisons are made between expected and observed adjustments. Structural tuning uses beliefs specified about possible improvements to our simulator and sampling approaches to specify the necessary beliefs for tuning. For example, specifying the bounds of a prior non-implausible set for the tuning parameters or a second order specification for diagnostic quantities. As this is performed within the simulator, it is termed internal analysis. When observations are made, a sample from the actual non-implausible set of tuning parameters and diagnostic quantities are calculated. Finally, comparisons may be made between i) expected non-implausible sets and diagnostic quantities and ii) actual non-implausible sets and diagnostic quantities. This means that when tuning, we can assess whether the outputs from the tuning process are commensurate with our prior specifications. Then, even if a simulator can be tuned we can assess whether the tuning inputs and transfer function are resolving the discrepancies that we have considered and, if so, there is a basis to believe the predictive properties of the tuned simulator.

6.3.1 Validating the Direct Tuning Assumption

The first goal is to *a-priori* validate the tuning assumption. The idea is that if $\mathcal{F}$ and $\{\mathcal{F}^\star, x^\star, w^\star\}$ are known, finding a parameter x'_E that satisfies equation 6.2.1 would validate the tuning assumption. Although this is a case which is unlikely to actually happen, it exemplifies that it is possible to validate the assumption. The reified simulator is unlikely to be known so beliefs about it are specified and a probabilistic model (emulator) built. Samples may be made from this probabilistic model (for example, using Gaussian processes – see definition 2.2.5) and, if for all samples a tuning input can be found, the tuning assumptions are commensurate with the expert's beliefs about the reified simulator. If the system values are uncertain, these must be sampled as well and, if for every sample of the system values and reified emulator a tuning input can be found, the tuning assumptions are commensurate with the expert's beliefs. Therefore, although challenging, it is possible to validate the assumption.

The interest is not in general whether the direct tuning assumption is correct. The interest is in whether making the direct tuning assumption enables efficient real world

inference using a complex computer model in that the error induced by the assumption is small compared to other sources of uncertainty. For example, if the difference between the tuned simulator and the reified version is small compared to measurement error and the discrepancy in the reified emulator, assuming a tuning input exists is not going to have a large impact on the inferences made. Therefore, whether the assumption is true – in other words a tuning input and transfer function can be found which return an exact match – is replaced by the more pertinent question of whether it is an acceptable assumption. Instead of sampling the reified emulator and system values, a calibration training-set is sampled, given the sample from the reified emulator and system values. Then, for every sample of the system values and reified emulator, a sample from the set of tuning inputs and transfer functions, which cannot be differentiated from measurement error and the discrepancy in the reified emulator, is returned. Essentially for every sampled calibration training-set, a non-implausible set is returned. If the non-implausible set is non-empty for every sample, then it is expected that our current simulator can be tuned to its reified counterpart.

The process sequentially considers different sources of discrepancy and proposes improved simulators. It was designed sequentially because considering all improvements together is a daunting task and breaking it up was found to result in more careful consideration by experts. An emulator of the improved simulator is built using, for example, the structural reification approach or by evaluating the improved simulator. The process begins with the current simulator given by $\mathcal{F}(x_S, x_E)$ and a new simulator is proposed

$$\mathcal{F}'(x_S, x_E, v) \text{ such that } \mathcal{F}(x_S, x_E) = \mathcal{F}'(x_S, x_E, v_0)$$

with additional inputs v which are a subset of w. For a range of values from the prior non-implausible set of $\{x_E^\star, v^\star\}$, a calibration training-set is emulated (or simulated) at $\{x_S^{\star(i)}\}_{i=1:N}$ and the measurement error sampled. This effectively creates an emulated calibration training-set, given $\mathcal{F}^\star = \mathbb{E}[\mathcal{F}']$, for a range of system values. The set of sampled calibration training-sets is notated $\mathcal{Z} = \{Z^{(i)}\}_{i=1:s}$ where s is the number of samples from the prior non-implausible set of system values. For every calibration training-set in $\mathcal{Z}$, the posterior non-implausible set of tuning parameters x_E' and associated adjusted transfer functions are calculated. This process is summarised in Algorithm 15. A more complete method would sample from the distribution of $\mathcal{F}'|D^f$. This would provide a better uncertainty analysis. However, for simplicity, in this thesis the analysis will be performed using the adjusted expectation of the emulator. This is because for our example

Algorithm 15 Internal Structural Tuning

Inputs: $D^{\mathcal{F}}$, $\{x_S^{\star(j)}\}_{j=1:N}$, s

Outputs: $\{\chi_E^{\prime(i)}\}_{i=1:s}$, associated-transfer-functions

1: Build an emulator for improved simulator $\mathcal{F}'(x_S, x_E, v)$.
2: **for** i in $1:s$ **do**
3: Sample $(x_E^{\star(i)}, v^{\star(i)})$ from the prior distribution of $(x_E^{\star}, v^{\star})$.
4: **begin:** Emulate the calibration training-set $\{Z^{(i)}\}_{i=1:s}$.
5: **for** j in $1:N$ **do**
6: Calculate the adjusted expectation of $\mathcal{F}'(x_S^{\star(j)}, x_E^{\star(i)}, v^{\star(i)})$.
7: Make a sample from $\epsilon_{me} + \epsilon_{md}^{\star}$ notated $\epsilon_{me}^{\star(i,j)} + \epsilon_{md}^{\star(i,j)}$.
8: Combine the adjusted expectation and $\epsilon_{me}^{(i,j)} + \epsilon_{md}^{\star(i,j)}$ to create sampled observation $Z_{(j)}^{(i)}$.
9: **end for**
10: **end**
11: Using $Z^{(i)}$ as the observations, calculate a sample from the non-implausible set of tuning parameters (notated $\chi_E^{\prime(i)}$) and the set of associated adjusted transfer functions for every non-implausible point in $\chi_E^{\prime(i)}$ when tuning $\mathcal{F}$ to $\mathcal{F}'$.
12: **end for**

we were able to run the simulator enough times that the uncertainty in $\mathcal{F}'$ was *not large*. Including samples of $\mathcal{F}'|D^{\mathcal{F}}$, particularly if $\mathcal{F}'$ cannot be evaluated, would be an interesting extension to this work. The analysis is similar except that the sample means and variances calculated in the following section are also over samples from $\mathcal{F}'|D^{\mathcal{F}}$.

6.3.2 Internal Analysis: Specifying Beliefs about Quantities of Interest

The outputs from Algorithm 15 also enable the specification of beliefs about quantities, which are not operationally defined, and diagnostics which are useful for analysing the inference performed. For example, a prior non-implausible set of tuning inputs should at least contain all values in the non-implausible sets for every sampled calibration training-set. In the rest of this section, a number of quantities will be defined and a strategy for approximating their second order specification presented. The quantities may be directly used in the inference or as diagnostics to ascertain whether the output of the inference may be taken at face value. They are not intended as an exhaustive set but as examples

of the types of calculations that may be performed in order to build confidence in the approach.

1. The quantities used directly in the inference are as follows.

   (a) The prior non-implausible set for the tuning parameters notated χ_E.

   (b) The second order specification for the transfer function Δ.

2. The diagnostics are as follows.

   (a) The difference between the expected system values and a summary statistic, see Section 5.2.1, of a sample from the non-implausible set of tuning parameters. This is notated $\bar{x}'_E - x^\star_E$ where $\bar{x}'_E$ is a summary statistic from the sample from the posterior non-implausible set of tuning parameters.

   (b) The expected resolved variance in the non-implausible set of tuning parameters, notated $\mathbb{RV}ar[\chi'_E]$.

   (c) The set $\{\bar{x}'_E, x^\star_E, v^\star\}$ where $v^\star$ is an additional input to an improved simulator.

   (d) The expected difference between the adjusted expectation of the reified simulator and an observation, notated $\mathbb{E}_{\{D^{\mathcal{F}}, D^f\}}[\mathcal{F}^\star(x^\star_S, x^\star_E, w^\star)] - z$.

A number of such calculations will now be discussed. To enable a more concise notation, we denote the sample mean and variance as $\mathcal{SM}$ and $\mathcal{SV}$. For example, $\mathcal{SM}(\chi_E^{\prime(i)})$ is the sample mean of every vector in $\chi_E^{\prime(i)}$, and $\mathcal{SV}(\chi_E^{\prime(i)})$ is the sample covariance matrix. First, how to calculate approximations for the quantities necessary for tuning is presented.

1.(a) A specification for the prior non-implausible set of tuning parameters can be specified using

$$\chi_E = \mathcal{CH}\left(\left\{\chi_E^{\prime(i)}\right\}_{i=1:s}\right) \tag{6.3.1}$$

where $\mathcal{CH}$ denotes the convex hull and $\chi_E^{\prime(i)}$ is a sample from the posterior non-implausible set of tuning parameters for calibration training-set $Z^{(i)}$.

1.(b) The second order specification for the regression parameters of the transfer functions

can be approximated using

$$\mathbb{E}\left[\beta^{\Delta}; \mathcal{F}^{\star} = \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}']\right] \approx \mathcal{SM}\left(\left\{\mathcal{SM}\left(\left\{\beta^{\Delta(i,j)}\right\}_{j=1:|\chi_E^{\prime(i)}|}\right)\right\}_{i=1:s}\right),$$

$$\mathbb{V}ar\left[\beta^{\Delta}; \mathcal{F}^{\star} = \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}']\right] \approx \mathcal{SV}\left(\left\{\mathcal{SM}\left(\left\{\beta^{\Delta(i,j)}\right\}_{j=1:\left|\chi_E^{\prime(i)}\right|}\right)\right\}_{i=1:s}\right) + \mathcal{SM}\left(\left\{\mathcal{SV}\left(\left\{\beta^{\Delta(i,j)}\right\}_{j=1:\left|\chi_E^{\prime(i)}\right|}\right)\right\}_{i=1:s}\right)$$

where $\beta^{\Delta(i,j)}$ is the GLS (see key-formula 2.2.1) approximated adjusted expectation of the regression component of the transfer function given the sampled calibration training-set $Z^{(i)}$ and the j^{th} non-implausible point in $\chi_E^{\prime(i)}$. The second order specification for the residual process parameters of transfer functions can be approximated using

$$\mathbb{E}\left[\mathcal{W}^{\Delta}(\mathcal{F}(x_S, x'_E)); \mathcal{F}^{\star} = \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}']\right] \approx$$
$$\mathcal{SM}\left(\left\{\mathcal{SM}\left(\left\{\left(\mathcal{W}^{\Delta(i,j)}(\mathcal{F}(x_S^{\star(1)}, x_E^{\prime(j)})), \cdots, \mathcal{W}^{\Delta(i,j)}(\mathcal{F}(x_S^{\star(N)}, x_E^{\prime(j)}))\right)\right\}_{j=1:|\chi_E^{\prime(i)}|}\right)\right\}_{i=1:s}\right),$$
$$\mathbb{V}ar\left[\mathcal{W}^{\Delta}(\mathcal{F}(x_S, x_E^{\prime(j)})); \mathcal{F}^{\star} = \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}']\right] \approx$$
$$\mathcal{SM}\left(\left\{\mathcal{SV}\left(\left\{\left(\mathcal{W}^{\Delta(i,j)}(\mathcal{F}(x_S^{\star(1)}, x_E^{\prime(j)})), \cdots, \mathcal{W}^{\Delta(i,j)}(\mathcal{F}(x_S^{\star(N)}, x_E^{\prime(j)}))\right)\right\}_{j=1:|\chi_E^{\prime(i)}|}\right)\right\}_{i=1:s}\right) +$$
$$\mathcal{SV}\left(\left\{\mathcal{SM}\left(\left\{\left(\mathcal{W}^{\Delta(i,j)}(\mathcal{F}(x_S^{\star(1)}, x_E^{\prime(j)})), \cdots, \mathcal{W}^{\Delta(i,j)}(\mathcal{F}(x_S^{\star(N)}, x_E^{\prime(j)}))\right)\right\}_{j=1:|\chi_E^{\prime(i)}|}\right)\right\}_{i=1:s}\right)$$

where $\left(\mathcal{W}^{\Delta(i,j)}(\mathcal{F}(x_S^{\star(1)}, x_E^{\prime(j)}), \cdots, \mathcal{W}^{\Delta(i,j)}(\mathcal{F}(x_S^{\star(N)}, x_E^{\prime(j)}))\right)$ is a vector of GLS approximated adjusted residual processes given sampled calibration training-set $Z^{(i)}$ and the j^{th} non-implausible point in $\chi_E^{\prime(i)}$. Although these equations above look challenging they are simply calculating expectations and variances over sets of sets which is not unusual in statistical analysis.

For the diagnostic quantities, the calculations are as follows.

2.(a) The second order specification of $\bar{x}'_E - x^{\star}_E$ is approximated using

$$\mathbb{E}\left[\bar{x}'_E - x^{\star}_E; \mathcal{F}^{\star} = \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}']\right] \approx \mathcal{SM}\left(\left\{\bar{x}'_E{}^{(i)} - x_E^{\star(i)}\right\}_{i=1:s}\right) \text{ and}$$
$$\mathbb{V}ar\left[\bar{x}'_E - x^{\star}_E; \mathcal{F}^{\star} = \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}']\right] \approx \mathcal{SV}\left(\left\{\bar{x}'_E{}^{(i)} - x_E^{\star(i)}\right\}_{i=1:s}\right).$$

2.(b) The second order specification of $\mathbb{R}\mathbb{V}ar[\chi'_E]$ is approximated using

$$\mathbb{E}\left[\mathbb{R}\mathbb{V}ar[\chi'_E]; \mathcal{F}^{\star} = \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}']\right] \approx \mathcal{SM}\left(\left\{\mathcal{SV}(\chi_E) - \mathcal{SV}(\chi_E^{\prime(i)})\right\}_{i=1:s}\right) \text{ and}$$
$$\mathbb{V}ar\left[\mathbb{R}\mathbb{V}ar[\chi'_E]; \mathcal{F}^{\star} = \mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}']\right] \approx \mathcal{SV}\left(\left\{\mathcal{SV}(\chi_E) - \mathcal{SV}(\chi_E^{\prime(i)})\right\}_{i=1:s}\right).$$

2.(c) The second order specification of $(\bar{x}'_E, x^\star_E, v^\star)$ is approximated using

$$\mathbb{E}\left[(\bar{x}'_E, x^\star_E, v^\star)\mathcal{F}^\star = \mathbb{E}_{D^\mathcal{F}}[\mathcal{F}']\right] \approx \mathcal{SM}\left(\left\{\left(\bar{x}'_E{}^{(i)}, x^{\star(i)}_E, v^{\star(i)}\right)\right\}_{i=1:s}\right) \text{ and}$$
$$\mathbb{V}ar\left[(\bar{x}'_E, x^\star_E, v^\star)\mathcal{F}^\star = \mathbb{E}_{D^\mathcal{F}}[\mathcal{F}']\right] \approx \mathcal{SV}\left(\left\{\left(\bar{x}'_E{}^{(i)}, x^{\star(i)}_E, v^{\star(i)}\right)\right\}_{i=1:s}\right).$$

2.(d) Given that $\mathcal{F}$ is tuned and not $\mathcal{F}'$, how much additional discrepancy is expected? This is approximated by

$$\mathbb{E}[\mathbb{E}_{\{D^\mathcal{F}, D^f\}}[\mathcal{F}^\star(x^\star_S, x^\star_E, w^\star)] - z; \mathcal{F}^\star = \mathbb{E}_{D^\mathcal{F}}[\mathcal{F}']] \approx$$
$$\mathcal{SM}\left(\left\{\mathcal{SM}\left(\left\{\mathbb{E}_{D^\mathcal{F}}[\mathcal{F}(x^{\star(j)}_S, \bar{x}'_E{}^{(i)})] + \mathbb{E}_{\{D^\mathcal{F}, Z^{(i)}\}}[\Delta(\mathcal{F}(x^{\star(j)}_S, \bar{x}'_E{}^{(i)}))] - Z^{(i)}_{(j)}\right\}_{i=1:s}\right)\right\}_{j=1:N}\right)$$
$$\mathbb{V}ar[\mathbb{E}_{\{D^\mathcal{F}, D^f\}}[\mathcal{F}^\star(x^\star_S, x^\star_E, w^\star)] - z; \mathcal{F}^\star = \mathbb{E}_{D^\mathcal{F}}[\mathcal{F}']] \approx$$
$$\mathcal{SM}\left(\left\{\mathcal{SV}\left(\left\{\mathbb{E}_{D^\mathcal{F}}[\mathcal{F}(x^{\star(j)}_S, \bar{x}'_E{}^{(i)})] + \mathbb{E}_{\{D^\mathcal{F}, Z^{(i)}\}}[\Delta(\mathcal{F}(x^{\star(j)}_S, \bar{x}'_E{}^{(i)}))] - Z^{(i)}_{(j)}\right\}_{i=1:s}\right)\right\}_{j=1:N}\right)$$
$$+ \mathcal{SV}\left(\left\{\mathcal{SM}\left(\left\{\mathbb{E}_{D^\mathcal{F}}[\mathcal{F}(x^{\star(j)}_S, \bar{x}'_E{}^{(i)})] + \mathbb{E}_{\{D^\mathcal{F}, Z^{(i)}\}}[\Delta(\mathcal{F}(x^{\star(j)}_S, \bar{x}'_E{}^{(i)}))] - Z^{(i)}_{(j)}\right\}_{i=1:s}\right)\right\}_{j=1:N}\right).$$

The diagnostic quantities may be calculated in the external analysis using the actual observed calibration training-set. Then using these second order specifications, we may assess whether, given our prior beliefs about the improved simulator, the revealed quantities are larger or smaller than expected.

If $\mathcal{F}'$ is not reified, a further simulator $\mathcal{F}''$, which is nested in $\mathcal{F}'$, is proposed. Then, if $\mathcal{F}$ can be tuned to $\mathcal{F}'$, we attempt to tune $\mathcal{F}$ to $\mathcal{F}''$. Then, if $\mathcal{F}$ cannot be tuned to $\mathcal{F}''$, we attempt to tune $\mathcal{F}'$ to $\mathcal{F}''$. This process continues until an expert asserts that there are no further improvements – in other words, their reified simulator has been considered – or any further improvements are small enough that it is likely that the current simulator may be tuned to the reified case. This might be the case when any further improvements are expected to be similar to the previous improvements but have a smaller unidentifiable impact on the output.

This process enables the specification of beliefs about quantities which are not operationally defined using the same level of elicitation as structural reification. The outputs are also useful for decision making as, if the current simulator available is not tunable, it would not be worth making the calibration observations as the goals of the study cannot be met, using this analysis, until the simulator is improved. Then, a choice may be made as to which improvement should be made to give the best chance of creating a tunable simulator. For example, the improvement which tunes out the discrepancy induced by the most subsequent improvements.

External Analysis

Once internal analysis has been performed, an external analysis is performed using actual observations. Then a set of external measures may be utilised to compare the expected inferred quantities against actual inferred quantities. Using the actual calibration training we calculate, i) a sample from the actual non-implausible set of tuning parameters, ii) the associated adjusted beliefs about the transfer functions and iii) the diagnostic quantities. Using the second order specifications calculated in the previous section, a number of diagnostics may be calculated for which large values indicate possible cause for concern.

1. The first diagnostic compares the revealed and expected difference between a summary statistic of the non-implausible set and the expected system values. It is given by

$$(\bar{x}'_E - \mathbb{E}[x^\star_E] - \mathbb{E}\left[\bar{x}'_E - x^\star_E; \mathcal{F}^\star = \mathbb{E}_{D^\mathcal{F}}[\mathcal{F}^\star]\right])^T \left(\mathbb{V}ar[x^\star_E] + \mathbb{V}ar\left[\bar{x}'_E - x^\star_E; \mathcal{F}^\star = \mathbb{E}_{D^\mathcal{F}}[\mathcal{F}^\star]\right]\right)^{-1}$$
$$(\bar{x}'_E - \mathbb{E}[x^\star_E] - \mathbb{E}\left[\bar{x}'_E - x^\star_E; \mathcal{F}^\star = \mathbb{E}_{D^\mathcal{F}}[\mathcal{F}^\star]\right])$$

   where $\bar{x}'_E$ is the actual summary statistic of the non-implausible set given the observed calibration training-set. Large numbers indicate that the tuned version is further away than expected meaning that the tuning process is resolving additional discrepancy which has not been considered.

2. The second measure is the discrepancy in the adjusted expectation of the system values (adjusted by the summary statistic of the sample from the non-implausible set of tuning parameters).

$$(\mathbb{E}_{\bar{x}'_E}[(x^\star_E, w^\star)] - \mathbb{E}[(x^\star_E, w^\star)])^T (\mathbb{R}\mathbb{V}ar_{\bar{x}'_E}[(x^\star_E, w^\star)])^{-1} (\mathbb{E}_{\bar{x}'_E}[(x^\star_E, w^\star)] - \mathbb{E}[(x^\star_E, w^\star)])$$

   where $\bar{x}'_E$ is the actual summary statistic of the non-implausible set given the observed calibration training-set. Here again large values indicate cause for concern.

3. The next measure compares the expected and actual resolution of the non-implausible set. This is given by

$$(\mathcal{SV}(\chi_E) - \mathcal{SV}(\chi'_E) - \mathbb{E}[\mathbb{R}\mathbb{V}ar[\chi'_E]; \mathcal{F}^\star = \mathbb{E}_{D^\mathcal{F}}[\mathcal{F}^\star]])^T (\mathbb{V}ar[\mathbb{R}\mathbb{V}ar[\chi'_E]; \mathcal{F}^\star = \mathbb{E}_{D^\mathcal{F}}[\mathcal{F}^\star]])^{-1}$$
$$(\mathcal{SV}(\chi_E) - \mathcal{SV}(\chi'_E) - \mathbb{E}[\mathbb{R}\mathbb{V}ar[\chi'_E]; \mathcal{F}^\star = \mathbb{E}_{D^\mathcal{F}}[\mathcal{F}^\star]])$$

   where χ'_E is a sample from the non-implausible set of tuning parameters calculated with the actual calibration observations. If the non-implausible set is much smaller than expected, there is evidence to suggest that there is some discrepancy between the actual simulator and the reified counterpart which cannot be resolved using

the tuning process. If the set is much larger than expected, it indicates that the difference between $\mathcal{F}$ and $\mathcal{F}^\star$ is smaller than specified. For example the specification of the improved simulator has specified the variation induced by w in $\mathcal{F}^\star$ as more unidentifiable than it is.

4. The final measure compares the expected and revealed remaining discrepancy. This is given by

$$(\mathbb{E}_{\{D^{\mathcal{F}},D^f\}}[\mathcal{F}^\star(x_S^\star,x_E^\star,w^\star)]]-z-\mathbb{E}[\mathbb{E}_{\{D^{\mathcal{F}},D^f\}}[\mathcal{F}^\star(x_S^\star,x_E^\star,w^\star)]-z;\mathcal{F}^\star=\mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}^\star]])^T$$
$$(\mathbb{V}ar[\mathbb{E}_{\{D^{\mathcal{F}},D^f\}}[\mathcal{F}^\star(x_S^\star,x_E^\star,w^\star)]-z;\mathcal{F}^\star=\mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}^\star]])^{-1}$$
$$(\mathbb{E}_{\{D^{\mathcal{F}},D^f\}}[\mathcal{F}^\star(x_S^\star,x_E^\star,w^\star)]]-z-\mathbb{E}[\mathbb{E}_{\{D^{\mathcal{F}},D^f\}}[\mathcal{F}^\star(x_S^\star,x_E^\star,w^\star)]-z;\mathcal{F}^\star=\mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}^\star]]).$$

Reified Bayesian analysis gives a coherent method for making inferences with complex computer models with belief statements about well defined quantities. Structural tuning gives confidence that tuning is reasonable as it validates – using the same statements about random quantities as structural reification – that it is unlikely that the simulator available will not give an acceptable match to the reified counterpart. The methodology we have presented suggested diagnostic measures and a process for obtaining an approximation of their second order specification. These measures and their associated specification may be used to ascertain – even if a non-implausible point is found – if there is evidence to suggest there is discrepancy between an expert's beliefs about the non-implausible sets and the actual non-implausible sets returned.

6.4 Reifying GEANT4 (G4)

It is now time to make inferences in the real world using G4. The rest of this chapter is a summation of the methodology set out in chapters 4 , 5 and this chapter. The hierarchical-superposition emulator presented in Section 4.3.4 will be used to make inferences for a reified simulator. The emulator of the reified simulator and the iterative history matching approach presented in Section 5.6.1 will be used to make inferences for the scatter present in an X-ray image and the composition of the entity being examined.

The first task is to identify sources of discrepancy.

1. The simulator available was not flexible enough to simulate entities of the required complexity.

2. A number parameters are hard coded. These parameters are correlated with equipment dependent system values, but how the parameters relate to the simulator inputs is uncertain.

3. The resolution of the inputs associated with the equipment dependent parameters is insufficient. In a cone beam simulation, every photon is sampled from the same spectrum; however, in reality, the filtration at the source will vary. This means that the spectrum of energies that photons are sampled from should depend on the initial direction the photon is travelling in.

4. The filtration at the source is encoded by a thickness of aluminium. This was chosen because it is the effective filtration of aluminium that is reported by manufacturers. Effective here means that it is not actually a piece of aluminium but an alloy of a number of different materials. Therefore, the fact that a thickness of aluminium is used as a proxy for a different alloy is likely to cause discrepancy.

5. The scintillation physics are not included. In digital X-ray examination an X-ray is absorbed by a scintillator which in turn produces visible light photons which are recorded by the detector. This G4 build only measures the energy deposited in the scintillator not the energy deposited at the digital detector. The reason for this is that producing and tracking the visible light photons would make a simulation prohibitively more expensive.

6. The detectors response to light photons is unlikely to be spatially invariant as different camera cards have different responses. This is corrected using *ad hoc* calibration but the correction is not perfect and will constitute another difference between GEANT4 and reality.

7. G4 has materials present coded by density and chemical composition. The material formula does not necessarily match the chemical composition and density of the physical material. For simple materials like aluminium, the discrepancy is expected to be small but for more complex alloys like cortical bone, it can be much larger.

6.4.1 The Simulator

The relationship between the simulator available and a more complex simulator which can model complex entities has already been proposed in Chapter 4 using a hierarchical-

superposition emulator. The simulator in this section will be assumed to be the output of the non-isotropic scatter emulator given by, $\mathcal{F}'(x_S, M = 1)$ (as defined in equation 4.3.5) and $\mathcal{F}'(x_S, M = 2)$ (as defined in equation 4.3.10). In this chapter, inputs which have been fixed so far will be released in order to tune to the reified case. The simulator to be tuned will be notated

$$\mathcal{F}'(x_S, x_E, M = 0) = \mathcal{F}'(x_S, x_E, M = 1) + \mathcal{F}'(x_S, x_E, M = 2).$$

Here x_E are parameters controlling the kV and the thickness of a cuboid of aluminium directly in front of the source, referred to as filtration. As discussed in Section 3.5.1, changing kV and adding filtration are two methods of altering the spectrum produced by an X-ray source. Therefore, these are two parameters which control the spectrum from which photons are sampled in pencil beam simulations. The emulator built in Section 4.3.4 was of a simulator where kV was fixed to 80 and 0mm of aluminium filtration. x_S is defined as in equation 4.3.1 and is a quantity of direct interest.

6.4.2 The Calibration Training-set

For the calibration training observations, a set of PMMA and aluminium cuboids were used. These were chosen because they are relatively cheap, their chemical composition and density is well known and they can be accurately manufactured to specifications. The source was collimated so that only the cuboids were illuminated by X-rays.

Definition 6.4.1. ***Collimator.*** *A piece of lead with a aperture in the middle, placed in front of the source, which can be adjust to control the region of the detector that is illuminated by X-rays.*

The set of calibration observations is notated $\{\mathcal{Z} = z^{(i,j)}(t^{(k)})\}_{i=1:2,j=1:9,k=1:\kappa\theta}$ where $i = 1$ is aluminium, $i = 2$ is PMMA and $\kappa\theta$ is the number of pixels illuminated on the detector. The total number of observations N is, therefore, 18. The 9 thicknesses of the cuboids were chosen so to span the prior range of possible intensities which may occur in clinical X-ray examination. Then for both materials and every pixel on the detector, the relationship between thickness and the observations is specified using the model

$$log(z^{(i,j)}((t))) = h(\eta^{(i,j)}(t))\beta^{(i)}(t) \oplus \epsilon^{(i,j)}(t) = (1 + \eta^{(i,j)}(t) + \eta^{(i,j)(t)2})\beta^{(i)}(t) \oplus \epsilon^{(i,j)}(t). \tag{6.4.1}$$

The residuals ϵ were assumed i) to be expectation zero, ii) to have constant variance and iii) $\mathbb{C}ov[\epsilon^{(i,j)}(t), \epsilon^{(i,j')}(t)] = 0$. The analysis is performed pixel wise so $\mathbb{C}ov[\epsilon^{(i,j)}(t), \epsilon^{(i',j')}(t')]$

was not considered. Similarly, the spatial correlation in the regression parameters was not considered. The log transformation linearises the data as the original form is well modelled by a sum over exponential functions (see equation 3.5.2 and discussion in Section 4.3.3). The set of $N_\eta = 50$ thicknesses were uniformly sampled within the convex hull of both thickness sets. Then the approximated adjusted expectation of the observed intensity is calculated for the 100 thickness material combinations using GLS estimates (see key-formula 2.2.1). For the variance of the measurement error and the reified model discrepancy, the quantity $\mathbb{V}ar[\epsilon_{me}^{(i,j)}(t)+\epsilon_{md}^{\star(i,j)}(t)] = h(\eta^{(i,j)}(t))\mathbb{V}ar[\beta(t)]h(\eta^{(i,j)}(t))^T+\mathbb{V}ar[\epsilon^{(i,j)}(t)]$ was used. This was chosen because the aim is to match the population mean of a stochastic simulator against the population mean of the observations. The model in equation 6.4.1 induces discrepancy as the relationship is not quadratic, therefore, an upper bound for the extent to which the reified simulator can mimic the observations is limited to the discrepancy in the quadratic model. The experiment was run at 80kV and variable current which is controlled by set-able milliamp seconds (mAs) in the tube. The mAs for each training sample was chosen so that every observation had approximately equal Poisson noise and they are notated $\{x_E^{f(i)}\}_{i=1:18}$. mAs are assumed to be proportional to the number of photons produced at the source. The images were then normalised to 1mAs assuming a linear relationship. This calibration stage meant that it is reasonable to assert that the residuals have constant variance. Given that the relationship is not quadratic, an alternative was to specify the residuals as correlated. However, testing showed that the difference in adjusted beliefs was *not large* so it was judged that the improvements gained by adding in a correlated residual process were not worth the pain-to-gain. The images were then sub-sampled (see equation 5.6.1 and accompanying discussion) by a factor 23 to enable calculations in a useful time frame. Factor 23 was chosen as it was the smallest factor for which the calculations could be performed without memory bottlenecks[2]. Figure 6.1 shows the adjusted regression model for a single pixel with a two sigma credible interval. This approach has been applied with many sets of training thicknesses on around six pieces of experimental equipment. The numbers stated in this section are for the example in this section. For other cases, they may vary. As such, a summary of the general method is presented in Algorithm 16.

[2]The major bottle neck was storing the look-up-table on dedicated graphics memory. My desktop had a GPU with 4GB of dedicated memory.

Algorithm 16 Create Reification Training-set

Inputs: $\{S^{(i)}\}_{i=1:N}$, $\{x_E^{f(i)}\}_{i=1:N}$, N_η

Outputs: D^f, $\{\mathbb{V}ar[\epsilon_{me}^{(i,j)} + \epsilon_{md}^{\star(i,j)}]\}_{i=1:2,j=1:\frac{N_\eta}{2},k=1:\kappa\theta}$

begin: Take measurements and calibrate.

for j in $1 : N$ **do**

  Set $z^{(j)} = \frac{f(S^{(j)}, x_E^{f(j)}) \oplus \epsilon_{me}^{(j)}}{x_E^{f(j)}}$.

end for

end

Reorganise $\{z^{(j)}\}_{j=1:N}$ into $\mathcal{Z} = \{z^{(i,j)}(t^{(k)})\}_{i=1:2,j=1:\frac{N}{2},k=1:\theta\kappa}$.

for k in $1 : \theta\kappa$ **do**

  for i in 1:2 **do**

    begin: Calculate GLS estimates for adjusted expectation and variance of i) $\beta^{(i)}(t^{(k)})$ and ii) measurement error and reified model discrepancy variance.

    for j in $1 : \frac{N_\eta}{2}$ **do**

      Set

$$D_{(i,j)}^f(t^{(k)}) = h(\eta^{(i,j)})\mathbb{E}_{\mathcal{Z}_{(i,:)}(t^{(k)})}[\beta^{(i)}(t^{(k)})]$$

and

$$\mathbb{V}ar[\epsilon_{me}^{(i,j)}(t^{(k)}) + \epsilon_{md}^{\star(i,j)}(t^{(k)})] = h(\eta^{(i,j)})\mathbb{V}ar_{\mathcal{Z}_{(i,:)}(t^{(k)})}[\beta^{(i)}(t^{(k)})]h(\eta^{(i,j)})^T + \hat{\sigma}_\epsilon.$$

    end for

    end

  end for

end for

6.4.3 Emulating GEANT4

Now the set of observations and measurement errors have been stated, the next step is to build the emulator of G4. To emulate over i) thickness, ii) material and iii) space, two multiple output emulators were used

$$\log(\mathcal{F}'(\eta, \mu = Al, x_E)) = \mathcal{M}^{Al}(\eta, x_E) \oplus \mathcal{W}^{Al}(\eta, x_E) \text{ and}$$
$$\log(\mathcal{F}'(\eta, \mu = PMMA, x_E)) = \mathcal{M}^{PMMA}(\eta, x_E) \oplus \mathcal{W}^{PMMA}(\eta, x_E).$$

The simulator here is the output of the hierarchical-superposition emulator with a cuboid entity of thickness η. Therefore, $\log(\mathcal{F}'((\eta, \mu = Al), x_E))$ is a vector with $\theta\kappa$ elements.

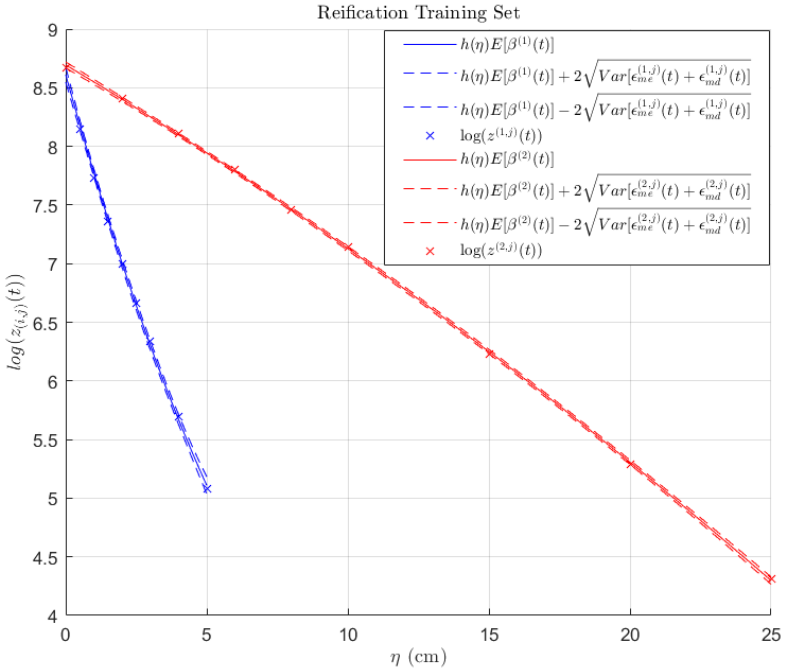

Figure 6.1: Tuning observations plotted against thickness for one pixel on the detector. Blue is aluminium and red is PMMA.

To adjust beliefs about the simulator output everywhere, pencil beam simulations with the same entities as the calibration training-set were run at a number of locations in the prior non-implausible set for x_E. Then, for each pencil beam simulation, the hierarchical-superposition emulator was evaluated assuming that $\mathcal{W}^a = 0$. This was considered reasonable as the geometry of the calibration entities means that the error between emulator and G4 is small, see Figure 4.7 and discussion in Section 4.3.4. A summary of the algorithm for creating an output from the hierarchical-superposition emulator using a single pencil beam is given in Algorithm 17. In this example, the tuning inputs were kV and filtration at the source. This was chosen using *ad hoc* analysis of the simulator which involved seeing if the simulator could be tuned, then expanding the prior non-implausible set of tuning parameters or adjusting other inputs to the simulator. This choice was subsequently supported by structural tuning analysis, see Appendix 6.3. The same vector of basis functions, namely $h(\eta, x_E) = (1, \eta, x_{E(1)}, x_{E(2)}, \eta^2, x_{E(1)}^2, x_{E(2)}^2)$, was considered appropriate for both the aluminium and PMMA emulators, and for all locations on the output field. The residual process covariance was specified as separable (see equation 2.2.7) using equation 2.2.8 with $p = 2$ and correlation lengths $l_3 = (1.5, 0.2, 1)$ for aluminium and $l_4 = (7.5, 0.2, 2)$ for PMMA. This form was chosen using validation diagnostics and graphical analysis. The

Algorithm 17 Predict $\mathcal{F}'$ for cuboid

Inputs: x_E, μ, η, l_1, a, b

Outputs: $\mathcal{F}'(\mathcal{G}(\mu,\eta,\mu,\eta), x_E, M=0)$

1: Evaluate simulator $\mathcal{F}(\mathcal{G}(\mu,\eta,\mu,\eta), x_E, M=t^c)$.

2: Radially collapse the simulator output to calculate F.

3: Using F, l_1, a, b and the emulator given by equation 4.3.11 and 4.3.12, build a look-up-table for every possible radial distance.

4: Use the look-up-table and pixel t^c to calculate scatter component $\mathcal{F}'(\mathcal{G}(\mu,\eta,\mu,\eta), x_E, M=0)$.

variance of the simulator output was specified directly (see equation 4.2.1 and discussion in Section 4.3.4) as proportional to the sample mean of the simulator output vector

$$\mathbb{V}ar[log(\mathcal{F}'_{(t)}(\eta,\mu,x_E))] = \frac{a_3}{\bar{\mathcal{F}}(\eta,\mu,x_E)} + b_3$$

where $\bar{\mathcal{F}}(\eta,\mu,x_E) = \frac{1}{\theta\kappa}\sum_i \mathcal{F}'_{(t^{(i)})}(\eta,\mu,x_E)$ is a scalar, $a_3 = 10$ and $b_3 = 0.0000000001$. This means that the covariance specification is separable (see equation 2.2.7) and, therefore, the GLS approximated adjusted expectations (see key-formula 2.2.1) were used to approximate adjusted beliefs. This simplification of the stochastic component was chosen using validation diagnostics and graphical analysis, see Section 2.2.2. It had significant computational advantages over a stochastic emulator with its population variance dependent on t and − as ensuring the algorithm could be evaluated in a useful time-frame, was a key consideration in this work − the computational advantages were considered worth the error in uncertainty estimation.

A lattice design was used for the kV and filtration. This was chosen as it was easy to implement and only integer values of kV could be input into this build of G4. 750 values were chosen in the range $(0\text{mm}, 70\text{kV}) \rightarrow (5\text{mm}, 90\text{kV})$. For this source, the prior expected system values were $\mathbb{E}[x^\star_E(t)] = (3, 80)$ so the range was large enough to be able to tune out the discrepancy but not so large that tuned values a *large* distance from the prior expected system values were acceptable.

In this example, a large number of simulator evaluations were available. In earlier versions of this method, a wave based approach was used which utilised many less simulations. However, it was decided that running the simulator online would not be a feasible option for robust software. Therefore, all simulations would have to be performed offline and, as long as the method is efficient, using all simulations when building the emulator

should give a similar result to a wave based approach. This also simplifies the analysis because the emulator uncertainty is small compared to other sources of uncertainty and can be ignored without a large impact on the inferences performed.

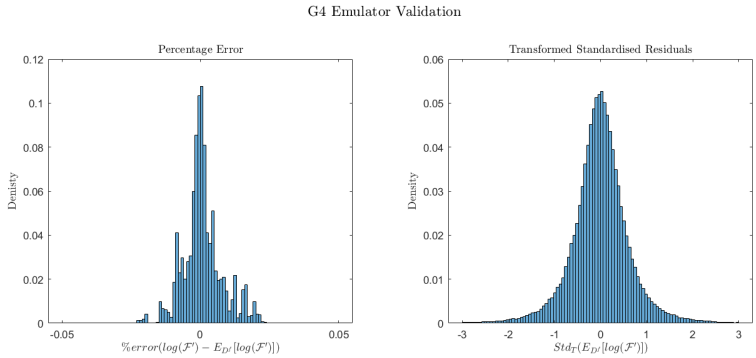

Figure 6.2: Left: percentage error histogram for validation points. Right: transformed standardised residual histogram for validation points.

To validate the emulator, a set of 100 points were sampled from the simulation design and used for validation. Figure 6.2 shows the percentage error for the validation points and the transformed standardised residuals, see definition 2.2.9. The percentage error is small validating that the error induced by the emulator is small compared to other sources of uncertainty. The percentage error distribution has some multi-modal features because the posterior correlation in the output remains. This correlation is well represented by the emulator which is shown by the uni-modal nature of the distribution of the transformed residuals. The uncertainty is slightly overestimated but overall the emulator performs well. The method for building the emulator of G4, for reifying G4, is summarised in Algorithm 18.

6.4.4 Tuning G4

Transfer Function

The next stage was to choose a model for the transfer function. The minimum requirement for feasible inference using this method is that the transfer function is independent of the simulator input given the simulator output. Therefore, it is assumed that equation 6.2.2 holds. The transfer function was specified as

$$\Delta_{(t)}(\log(\mathcal{F}_{(t)}(x_S^\star, x'_E))) = \sum_m \beta^{\Delta}_{(m,t)} h_{(m)}(\log(\mathcal{F}_{(t)}(x_S^\star, x'_E))) \oplus \mathcal{W}^{\Delta}_{(t)}(\log(\mathcal{F}_{(t)}(x_S^\star, x'_E)))$$

Algorithm 18 Build the emulator of G4

Inputs: $X_E^{\mathcal{F}}$, $\{\eta^{(i,j)}\}_{i=1:2,j=1:\frac{N}{2}}$, $\mu^{(1)}$, $\mu^{(2)}$, l_1, l_3, l_4, a, b, a_3, b_3

Outputs: G4-Emulator-M1, G4-Emulator-M2

begin: Create emulator training-set.
for k in 1 : nrow($X_E^{\mathcal{F}}$) **do**
  for i in 1 : 2 **do**
    for j in 1 : $\frac{N}{2}$ **do**
      Using $X_{E(k)}^{\mathcal{F}}$, $\eta^{(i,j)}$, $\mu^{(i)}$, l_1, a and b, apply Algorithm 17 to calculate approximate adjusted expectation of the simulated image, notated $F^{(i,j,k)}$.
    end for
  end for
end for
end
Using $F^{(1,:,:)}$, l_3, a_3 and b_3, build and validate the emulator of $\log(\mathcal{F}'(\eta, \mu = Al, x_E))$ to create G4-Emulator-M1.
Using $F^{(2,:,:)}$, l_4, a_3 and b_3, build and validate the emulator of $\log(\mathcal{F}'(\eta, \mu = PMMA, x_E))$ to create G4-Emulator-M2.

where $h(x) = (1, x, x^2)$. The residual process was specified as uncorrelated with constant variance and GLS estimates (see key-formula 2.2.1) were used as point estimates for the regression and variance hyper-parameters. The transfer functions were allowed to vary spatially for two reasons. The first is that the detectors response is unlikely to be spatially uniform. The second is that x_E is under-resolved in that the spectrum from which photons are sampled should vary spatially. Initially the residual process was specified as correlated. However, as experimental techniques and the simulator's bugs were sequentially removed, this became unnecessary and the uncorrelated version was significantly cheaper.

Inversion and Prediction

To find the posterior non-implausible set for the tuning parameters, the analysis was performed pixel wise. For each pixel, a sample from the posterior non-implausible set of x'_E and associated transfer functions are calculated. The full calibration data-set is split into a set of PMMA and aluminium and then beliefs about the transfer function are adjusted using each data-set separately. The implausibility measure is given by equation 6.2.4 and

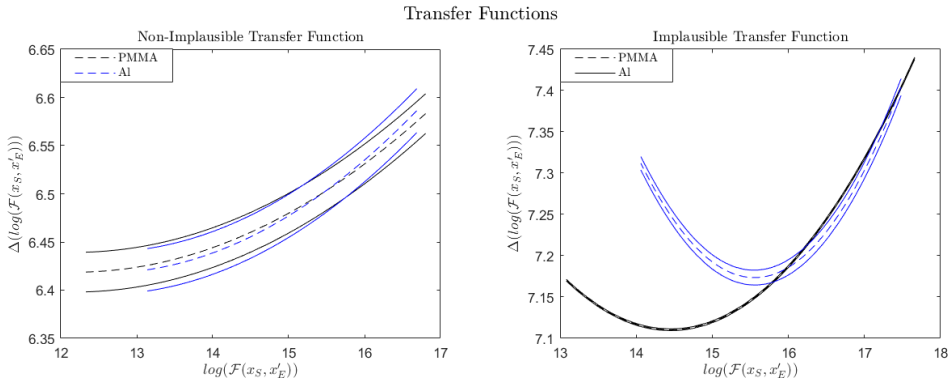

Figure 6.3: Left: adjusted beliefs about a transfer function for a non-implausible point. Right: adjusted beliefs about a transfer function for an implausible point.

any value over two was considered implausible. It was assumed in this version that the emulator uncertainty was zero. This was because, given the number of simulator training points, the emulator uncertainty is small compared to the other sources of uncertainty and testing showed that adding in emulator uncertainty had almost no impact on the non-implausible set returned. This assumption simplified matters as the uncertainty in the simulator did not need to be propagated through the transfer function. If a wave analysis is performed or the simulator uncertainty is larger, it may be propagated through to the output using the method outlined in Appendix A.7. A summary of the method for calculating a sample from the posterior non-implausible set of tuning inputs and associated transfer functions for a single pixel is given in Algorithm 19.

Figure 6.3 shows adjusted beliefs about the transfer functions for a non-implausible and an implausible point, for one pixel on the detector. On the left, the 2 sigma credible intervals overlap everywhere which indicates that the implausibility will never be greater than two. On the right, the credible intervals never overlap which is evidence that the difference between tuned simulator and the reified counterpart does not satisfy equation 6.2.2.

Figure 6.4 illustrates the inversion and prediction steps for a single pixel. The set of points which are considered tunable to the reified case have been significantly reduced and the system values $(80kV, 2mm)$ are not in the non-implausible set. This shows that if the discrepancy was not properly considered, evaluating the simulator at the *a-priori* system values would not have yielded an acceptable match. A major advantage of this approach, over other simulation based scatter removal methods, is that discrepancy is properly considered. The percentage error, for both integrated and the point estimate

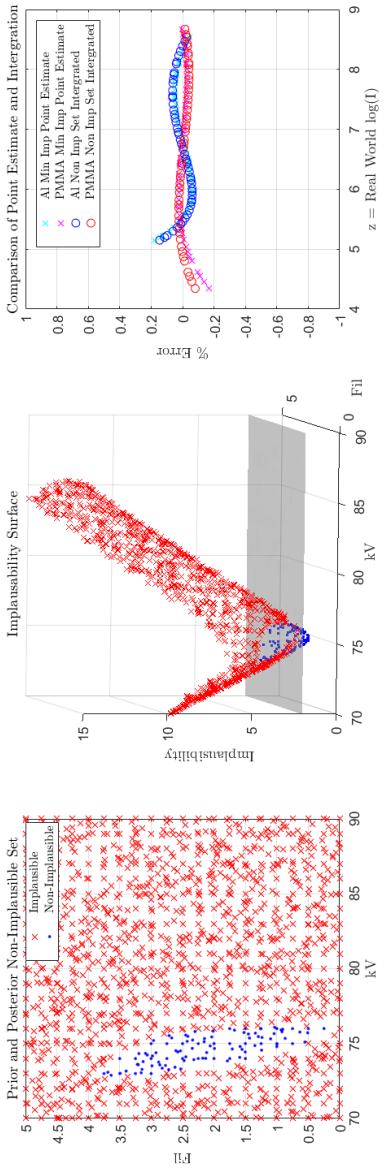

Figure 6.4: Top: percentage error between expected value of the reified emulator and the observations. Centre: implausibility surface. Bottom: prior and posterior non-implausible sets.

predictions, is small indicating no evidence to suggest an unacceptable match to the reified simulator. The next step is, given a non-implausible set for every pixel, how to perform prediction for the entire reified emulator output. In this work, a single kV and filtration

Algorithm 19 Calculate Non-implausible Tuning Inputs and Transfer Functions for a Single Pixel

Inputs: $D^f(t)$, G4-Emulator-M1, G4-Emulator-M2, $\mathcal{X}_E$, $\mathbb{V}ar[\epsilon_{me}(t) + \epsilon^{\star}_{md}(t)]$, ImThresh

Outputs:$\mathcal{X}^{\star}_E$, $\Delta^{\star}$.

1: Set $\mathcal{X}^{\star}_E = \mathcal{X}_E$
2: **for** i in 1:nrow($\mathcal{X}_E$) **do**
3: Using G4-Emulator-M1, $\mathbb{V}ar[\epsilon_{me}(t) + \epsilon^{\star}_{md}](t)]$ and material one training-set, adjust beliefs about transfer function regression parameters and variance of the residual process given $x'_E = \mathcal{X}_{E(i)}$, notated $\hat{\beta}^{\Delta(1,i)}$, $\mathbb{V}ar[\hat{\beta}^{\Delta(1,i)}]$ and $\hat{\sigma}^{\Delta(1,i)}$.
4: Using G4-Emulator-M2, $\mathbb{V}ar[\epsilon^{(i,j)}_{me}(t) + \epsilon^{(i,j)}_{md}(t)]$ and material two training-set, adjust beliefs about transfer function regression parameters and variance of the residual process given $x'_E = \mathcal{X}_{E(i)}$, notated $\hat{\beta}^{\Delta(2,i)}$, $\mathbb{V}ar[\hat{\beta}^{\Delta(2,i)}]$ and $\hat{\sigma}^{\Delta(2,i)}$.
5: Using $\mathbb{V}ar[\epsilon_{me}(t) + \epsilon^{\star}_{md}(t)]$, $\hat{\beta}^{\Delta(1,i)}$, $\hat{\sigma}^{\Delta(1,i)}$, $\mathbb{V}ar[\hat{\beta}^{\Delta(1,i)}]$, $\hat{\beta}^{\Delta(2,i)}$, $\mathbb{V}ar[\hat{\beta}^{\Delta(2,i)}]$ and $\hat{\sigma}^{\Delta(2,)}$, calculate implausibility $\mathcal{IM}^{(i)} = \mathcal{IM}(\mathcal{X}_{E(i)}, D^f(t))$.
6: **if** $\mathcal{IM}^{(i)} >$ ImThresh **then**
7: Remove $\mathcal{X}_{E(i)}$ from $\mathcal{X}^{\star}_E$.
8: **end if**
9: **if** $\mathcal{IM}^{(i)} <$ ImThresh **then**
10: Add $\hat{\beta}^{\Delta(i)}$ and $\hat{\sigma}^{\Delta(i)}$ to $\Delta^{\star}$.
11: **end if**
12: **end for**

is chosen and then the transfer functions are allowed to vary spatially. Point estimate prediction was chosen for two reasons. The first is that it is cheaper to not integrate over a set and computational expense was a prominent factor in our decisions. The second is that allowing kV and filtration to vary at the detector will induce another form of discrepancy. Allowing kV and filtration to vary at this stage would mean that the output at pixel (i, j) would be from a simulator which assumes the entire experiment has spectrum defined by $kV_{(i,j)}$ and $Fil_{(i,j)}$. In fact, the reified version should have a variable spectrum entering the entity which implies that a spatially variable x_E should be included in the hierarchical-superposition emulator from Chapter 4. Adding variable kV and filtration at the detector had a large cost and whether it is in-fact an improvement is an open research question. Adding it in the hierarchical-superposition emulator would be an improvement

and an interesting extension to this work. Therefore, a single kV and filtration was used for prediction which takes place in two steps: i) choose a single kV and filtration and ii) , given that kV and filtration, adjust beliefs about a transfer field. Allowing the transfer function

Algorithm 20 Calibrate G4 to Reified Simulator

Inputs: $\{S^{(i)}\}_{i=1:N}$, $\{x_E^{f(i)}\}_{i=1:N}$, N_η, $X_E^{\mathcal{F}}$, $\{\eta^{(i,j)}\}_{i=1:2,j=1:\frac{N}{2}}$, $\mu^{(1)}$, $\mu^{(2)}$, l_1, l_3, l_4, a, b, a_3, b_3 , $\mathcal{X}_E$, ImThresh

Outputs: $\hat{x_E}'$, $\hat{\beta}^{\Delta}$, σ_{md}

Create reification training-set: using $\{S^{(i)}\}_{i=1:N}$, $\{x_E^{f(i)}\}_{i=1:N}$, N_η, $X_E^{\mathcal{F}}$, $\{\eta^{(i,j)}\}_{i=1:2,j=1:\frac{N}{2}}$, $\mu^{(1)}$, $\mu^{(2)}$, apply Algorithm 16.

Build emulator of G4: using $X_E^{\mathcal{F}}$, $\{\eta^{(i,j)}\}_{i=1:2,j=1:\frac{N}{2}}$, $\mu^{(1)}$, $\mu^{(2)}$, l_1, l_3, l_4, a, b, a_3, b_3, apply Algorithm 18.

begin: Find non-implausible set for every pixel.

for i in $1:\theta\kappa$ **do**

    Using $D^f(t^{(i)})$, G4-Emulator-M1, G4-Emulator-M2 , $\mathcal{X}_E$, $\mathbb{V}ar[\epsilon_{me}^{(i,j)}(t^{(i)})+\epsilon_{md}^{(i,j)}(t^{(i)})]$ and ImThresh, apply Algorithm 19.

end for

begin:

Set $\hat{x_E}'$ as the point which is non-implausible for the most pixels.

for i in $1:\theta\kappa$ **do**

    Set $\hat{\beta}^{\Delta}(t^{(i)})$ as the adjusted expectation of $\beta^{\Delta}(t^{(i)})$ given $\hat{x_E}'$.

end for

for i in $1:2$ **do**

    for j in $1:\frac{N}{2}$ **do**

        Using $\mathcal{F}'(\mathcal{G}(\mu^{(i)},\eta^{(j)},\mu^{(i)},\eta^{(j)}), x_E = \hat{x_E}', M = 0)$ and $\beta^{\Delta}(t^{(i)})$, calculate the residual

$$e^{(i,j)} = z^{(i,j)} - \mathbb{E}_{\{D^f,D^{\mathcal{F}}\}}[\mathcal{F}^{\star}(\mathcal{G}(\mu^{(i)},\eta^{(j)},\mu^{(i)},\eta^{(j)}), x_E^{\star}, w^{\star})]$$

    end for

end for

for i in $1{:}\theta\kappa$ **do**

    Set $\sigma_{md}(t^{(i)})$ as the sample variance of $e^{(:,:)}(t^{(i)})$.

end for

to vary spatially will resolve some of the discrepancy induced by photons being sampled

from a single spectrum. The kV and filtration chosen was the point in the non-implausible set of the most pixels. In other words, it is the input for which the match to observed behaviour is acceptable at the most locations on the detector. The advantage of this approach is that it is simple. However, further validation is required. The implausibility, given the chosen tuning point, for every pixel should be checked to make sure that it is non-implausible at the majority of locations on the detector, see Figure 6.6. If there are regions on the detector which are implausible, a better choice may be the point which is not too implausible at every location. However, this situation has not occurred in any of the examples we have tested. A summary of the method for tuning G4 to its reified counterpart is given in Algorithm 20.

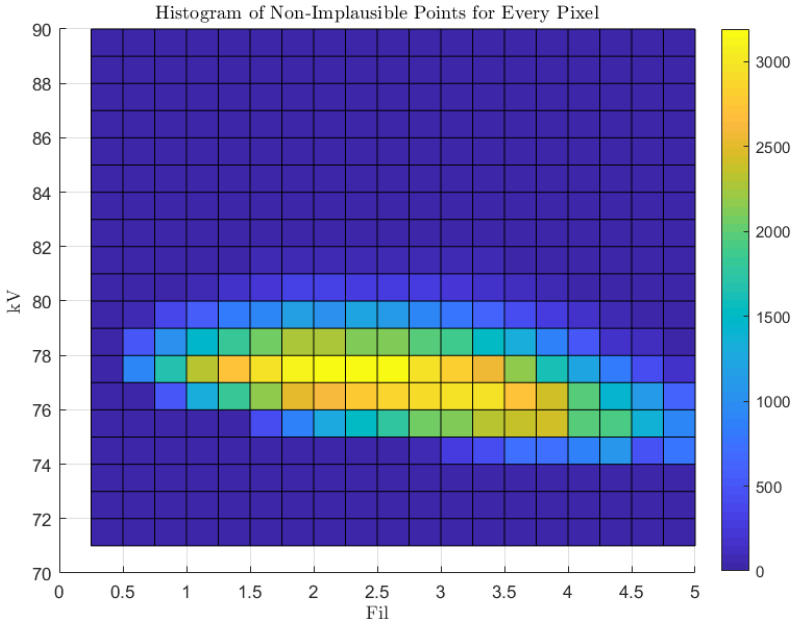

Figure 6.5: Heat map of a histogram of non-implausible points for every pixel on the detector.

Figure 6.5 shows a heat-map of a histogram of every non-implausible point for every pixel. Large portions of the space have been ruled out for every pixel. The point $x_E = (77, 2.5)$ was in the most non-implausible sets, 93% in total.

Figure 6.6 shows the implausibility for $\hat{x_E}' = (77, 2.5)$ for every pixel on the detector. The region where it is implausible is near the edge. This is not surprising as the collimation, see definition 6.4.1, in the real world is not modelled in the simulator. Collimation will change the spectrum of photons incident on the entity at the edges, making it extremely different from the rest of the image.

Figure 6.7 shows the percentage error between the adjusted expectation of the reified

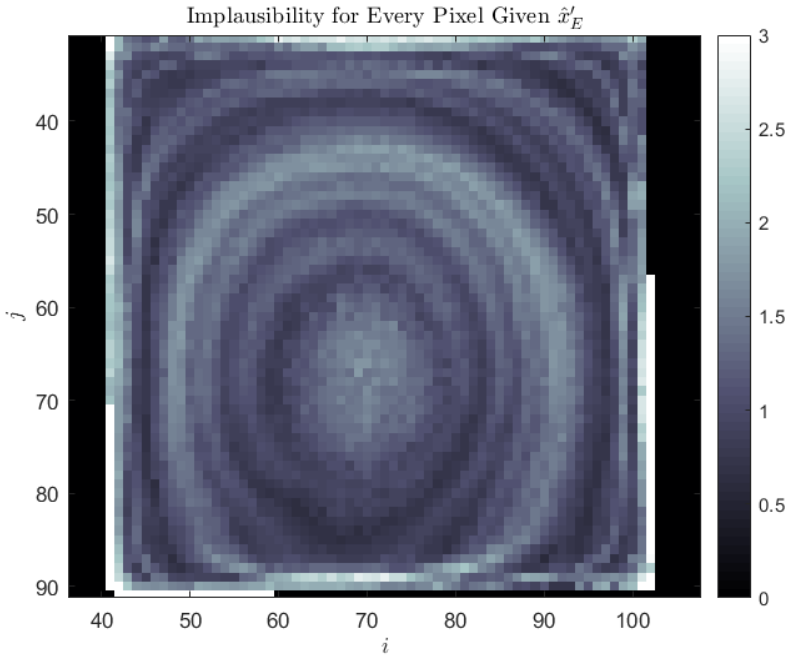

Figure 6.6: Implausibility for every pixel for the chosen tuning point. Black shows were the detector has not been illuminated by the direct beam.

emulator and the calibration observations. The error is mostly within $[\pm 5]\%$ which was considered accurate enough for the purpose of image improvement and entity composition inference. There is still spatial correlation in the errors, particularly towards edges of the thick samples.

Figure 6.8 shows the sample variance of the residual between the adjusted expectation of the reified emulator and the observation. This is a good point estimate for the uncertainty in the emulator, transfer function, measurement error and reified model discrepancy. This estimate is used in uncertainty judgements later in this chapter. Spatial correlation is present in the variance image which is partly due to the discrepancy which is induced by fixing filtration to a single value. As the analysis is performed pixel wise, the spatial correlation is not considered. The spatial correlation could be included in the analysis using, for example, the implausibility measure given in equation 5.2.2. However, as the variance of this measure is dependent on unspecified distributions, choosing a threshold is more challenging as Pukelsheim's 3 sigma rule can not be directly applied [47].

To validate that the transfer function is invariant to the entity dependent inputs, a set of stainless steel cuboids were used. A database of pencil beam simulations were performed at $(77, 2.5)$ where the entity was a cuboid of stainless steel and the thickness

was varied between 0 and 1000 microns. The adjusted expectation of $\mathcal{F}'$ was calculated using Algorithm 17. An emulator over thickness of stainless steel was built and used to find the posterior non-implausible thickness sets for stainless steel for the central pixel. The process is considered not invalidated if the known truth is in the non-implausible set, and a large portion of the set was declared implausible. Figure 6.9 shows the results of this validation procedure. The vee shaped black lines show the implausibility (equation 5.2.1) for one pixel on each observation for all prior non-implausible thicknesses. The posterior non-implausible sets are much smaller than the prior non-implausible set and always contains the known truth. This does not invalidate, at least for stainless steel, that the tuned simulator gives an acceptable match to the reified version.

This was considered an important proof point. Given a set of observations of i) PMMA and aluminium and ii) two simulated databases. Uncertainty in the thickness of stainless steel can be resolved. This is the solution to problem statement 1.1.4 and the solution is materially invariant. It is important to note that the material invariance assumes G4 models all materials equally well, however, further discussion on this is reserved for the cadaver study in Section 6.6.2.

In Appendix A.9, an illustrative example of the structural tuning approach is applied to G4. The simulator tuned in this section was finalised using an *ad-hoc* version of this technique. This essentially involved checking if it could be tuned and if it could not, i) considering the possible sources of discrepancy, ii) adjusting fixed parameters or iii) expanding the prior non-implausible space. It was inefficient and did not take into account all the information available. The structural tuning approach was created as a formal method of performing this type of analysis and, in hindsight, it would have saved a lot of time and effort.

6.5 Real World Scatter Removal and Composition Inference

Now all the tools to be able to perform real world composition inference and scatter correction have been developed. This section will combine the methods set out in the last three chapters into an algorithm which can provide clinical benefits. An emulator of a simulator which can model X-ray images of entities of required complexity is built using the hierarchical-superposition approach described in Chapter 4. The hierarchical-superposition emulator is related to its reified counterpart using the tuning approach presented in the previous section. Finally, the iterative history matching approach de-

scribed in Chapter 5 will be used to make inferences about the entity's composition and find a non-implausible point for predicting scatter.

To analyse the techniques, two studies were performed. The first uses quantitative phantoms to test against known ground truth as to whether non-implausible points may be used as point estimates for entity dependent system values. The second is a cadaver trial to test the image quality improvement on clinically relevant data. Assessing image quality is generally a difficult task. The metrics commonly used only tell part of the story and judgements are often subjective. The method we found most convincing and robust is visual comparison with equalised contrast stretch. The equalised contrast stretch looks at contrast-to-noise in large regions and allows assessment of overall image quality. The approach is to select a region-of-interest on all images that are being compared, normalise the regions to the same range and add a constant to the image so that a particular feature has on average equal intensity. For example, the region-of-interest could be a NOF and a particular feature could be the bone.

The most common form of scatter correction is an anti-scatter-grid (ASG) and this is considered current state-of-the-art in our analysis. We made this decision because anecdotal reports stated that the reason software solutions are rarely used is that they are not robust or simply provide lower quality images. This was confirmed by comparisons between our methods and other scatter correction techniques. However, due to commercial considerations, this analysis is not presented in this thesis so our comparisons are to a ASG.

A number of other metrics were also tested but are not presented here. The first was a signal-to-noise-ratio image which is calculated as the ratio of a moving average filtered image and a standard deviation filtered image. This is useful for assessing the local noise and the level of smoothing but it is not good at assessing diagnostic quality. Signal-to-noise in a region-of-interest is often used but we found it too dependent on the region chosen. Similarly, contrast-to-noise was too dependent on manually chosen regions. The method we performed shows the whole image and enables the reader to make their own assessment of image quality. The image quality has also be verified by third parties who were provided with the processed images and allowed to manipulate the images with their software of choice.

Method Summary and Proposed Algorithm

The different components of the analysis which have been presented need to be combined to perform an iterative history match and scatter removal in the real world. This combines the i) emulation strategy set out in Section 4.3.4, ii) iterative history matching approach set out in Section 5.6.1 and iii) reification strategy set out in Section 6.4. The observation is an X-ray image which may be decomposed into scatter and direct components

$$z = Y \oplus \epsilon_{me} = Y^{Dr} + Y^{Sc} \oplus \epsilon_{me}.$$

The relationship between the observation and the reified simulator is given by

$$Y = \mathcal{F}^{\star}(x_S^{\star}, x_E^{\star}, w^{\star}, M = 0) \oplus \epsilon_{md}^{\star}$$

where $\mathcal{F}^{\star}$ is the reified simulator of the system, $x_S^{\star}$ are the entity dependent system values, $(x_E^{\star}, w^{\star})$ are the equipment dependent system values and M is the mode of the simulator. The reified simulator is assumed to have three modes: direct, scatter and full.

$$\begin{aligned} Y^{Dr} &= \mathcal{F}^{\star}(x_S^{\star}, x_E^{\star}, w^{\star}, M = 1) \oplus \epsilon_{md}^{\star}(M = 1) \text{ and} \\ Y^{Sc} &= \mathcal{F}^{\star}(x_S^{\star}, x_E^{\star}, w^{\star}, M = 2) \oplus \epsilon_{md}^{\star}(M = 2) \text{ where} \\ \mathcal{F}^{\star}(x_S^{\star}, x_E^{\star}, w^{\star}, M = 0) &= \mathcal{F}^{\star}(x_S^{\star}, x_E^{\star}, w^{\star}, M = 1) + \mathcal{F}^{\star}(x_S^{\star}, x_E^{\star}, w^{\star}, M = 2) \\ \text{and } \epsilon_{md}^{\star} &= \epsilon_{md}^{\star}(M = 1) \oplus \epsilon_{md}^{\star}(M = 2). \end{aligned}$$

If the reified simulator was available, the method described in Section 5.6.1 could be applied. As the reified simulator is not available, a tuned simulator will be used instead. The simulator to be tuned, defined in Section 4.3.4, is given by

$$\mathcal{F}'(x_S, x_E, M = 0) = \mathcal{F}'(x_S, x_E, M = 1) + \mathcal{F}'(x_S, x_E, M = 2)$$

where inputs x_E are no longer fixed. The emulator of $\mathcal{F}'$ is built using a hierarchical-superposition strategy from Section 4.3.4. The relationship between $\mathcal{F}'$ and $\mathcal{F}^{\star}$ is given by

$$\log(\mathcal{F}^{\star}(x_S^{\star}, x_E^{\star}, w^{\star}, M = 0)) = \log(\mathcal{F}'(x_S^{\star}, x_E', M = 0)) + \Delta(\log(\mathcal{F}'(x_S^{\star}, x_E', M = 0))).$$

Here the form of $\Delta(\log(\mathcal{F}'(x_S^{\star}, x_E', M = 0)))$ is given by equation 6.4.1. The simulator available is informative for an improved one which is used to model entities of arbitrary complexity and a tuned version of that simulator is informative for a reified counterpart. This simulator is considered a direct simulator for reality. The offline part of the method

Algorithm 21 Scatter Removal Method: Offline Calculations

Inputs: $\{S^{G4(i)}\}_{i=1:\mathbb{N}_2}$, X, $X^{\mathcal{LU}}$, $\{S^{(i)}\}_{i=1:N}$, $\{x_E^{f(i)}\}_{i=1:N}$, N_η, $X_E^{\mathcal{F}}$, $\{\eta^{(i,j)}\}_{i=1:2,j=1:\frac{N}{2}}$, $\mu^{(1)}$, $\mu^{(2)}$, l_1, l_3, l_4, a, b, a_1, b_1, a_3, b_3, $\mathcal{X}_E$, ImThresh

Outputs: Direct-Beam-Emulator, $\mathcal{LU}$, $\hat{\beta}^{\Delta}$, σ_{md}

1: Using $\{S^{(i)}\}_{i=1:N}$, $\{x_E^{f(i)}\}_{i=1:N}$, N_η, $X_E^{\mathcal{F}}$, $\{\eta^{(i,j)}\}_{i=1:2,j=1:\frac{N}{2}}$, $\mu^{(1)}$, $\mu^{(2)}$, l_1, l_3, l_4, a, b, a_3, b_3 , $\mathcal{X}_E$ and ImThresh, apply Algorithm 20.

2: Using $\{S^{G4(i)}\}_{i=1:\mathbb{N}_2}$, a_1, b_1, l_3, X, $X^{\mathcal{LU}}$, a, b, l, l_1 and l_2, apply Algorithm 6 with the pencil simulations evaluate at $\hat{x_E}'$.

is given by Algorithm 21. This may be performed at the beginning of the day or a couple of times a day and should not affect clinical practice. For any x_S within the prior non-implausible set, adjusted beliefs about $\mathcal{F}^\star(x_S, x_E^\star, w^\star, M = 0)$ may now be calculated. This is performed in two simple stages: calculate the adjusted expectation for $\mathcal{F}'$ and run the adjusted expectation through the transfer function. For a new observation of an entity with unknown $x_S^\star$, the next step is to find a non-implausible point. If the exhaustive history match with random field priors is performed, no further assumptions are required. The problem is that the exhaustive history match is too expensive to return outputs in a useful time frame. As shown in Section 5.6.1 and summarised in Algorithm 9, the iterative approach can return a non-implausible point and improved image quality, but, requires an additional assumption.

Iterative History-Matching with Transfer Functions

Ideally we would like to assert

$$\log(\mathcal{F}^\star(x_S^\star, x_E^\star, w^\star, M = i)) = \log(\mathcal{F}'(x_S^\star, x_E', M = i)) \oplus \Delta(\log(\mathcal{F}'(x_S^\star, x_E', M = i))) \tag{6.5.1}$$

for $i = \{0, 1, 2\}$. However, Δ is not a linear map and, even if Δ is a linear map, this is making a strong assumption on the form of the discrepancy. It is saying non of the discrepancy can be attributed to only direct or scattering interactions. Given that scattering interactions are more complex than photoelectric absorption, it is in fact likely to be untrue. Therefore, even if Δ is a linear map, asserting 6.5.1 will induce an error.

Given that an error is induced, the next step is to assess the size of the error. The way in which the simulator is tuned implies that it is not large compared to the other sources

of uncertainty. Consider the differences

$$\Delta^1 = |\mathcal{F}^\star(x_S^\star, x_E^\star, w^\star, M=1) - \mathcal{F}'(x_S^\star, x_E'^\star, M=1)| \text{ and}$$
$$\Delta^2 = |\mathcal{F}^\star(x_S^\star, x_E^\star, w^\star, M=2) - \mathcal{F}'(x_S^\star, x_E'^\star, M=2)|.$$

If Δ^1 is much larger than Δ^2, the discrepancy is highly dependent on the proportion of the output induced by photons which have undergone a scattering interaction. The calibration training-set contains entities made of high scattering PMMA and low scattering aluminium. Therefore, if the discrepancy is mostly associated with scatter, the difference between the simulator and reified counterpart should be dependent on whether it is trained using PMMA or aluminium. However, the tuning procedure is designed to return points for which there is no evidence to suggest that the difference is dependent on whether PMMA or aluminium is used. This implies that if the point is non-implausible, the error induced is small enough to give an acceptable match.

If this argument is accepted, the problem is that Δ is not a linear map. To overcome this problem a correction is added to the scattering component

$$\mathcal{F}^\star(x_S^\star, x_E^\star, w^\star, M=2) = \exp(\log(\mathcal{F}'(x_S^\star, x_E'^\star, M=2)) + \Delta(\log(\mathcal{F}'(x_S^\star, x_E'^\star, M=2)))) + C \tag{6.5.2}$$

where

$$\begin{aligned} C = &\exp\left(\log\left(\mathcal{F}'\left(x_S^\star, x_E'^\star, M=0\right)\right) + \Delta\left(\log\left(\mathcal{F}'\left(x_S^\star, x_E'^\star, M=0\right)\right)\right)\right) \\ &- \exp\left(\log\left(\mathcal{F}'\left(x_S^\star, x_E'^\star, M=1\right)\right) + \Delta\left(\log\left(\mathcal{F}'\left(x_S^\star, x_E'^\star, M=1\right)\right)\right)\right) \\ &- \exp\left(\log\left(\mathcal{F}'\left(x_S^\star, x_E'^\star, M=2\right)\right) + \Delta\left(\log\left(\mathcal{F}'\left(x_S^\star, x_E'^\star, M=2\right)\right)\right)\right). \end{aligned} \tag{6.5.3}$$

This was chosen so that when expected scatter is removed from the observation, the expected value of what remains is

$$\exp\left(\log\left(\mathcal{F}'\left(x_S^\star, x_E'^\star, M=1\right)\right) + \Delta\left(\log\left(\mathcal{F}'\left(x_S^\star, x_E'^\star, M=1\right)\right)\right)\right).$$

This means the history matching step of the iterative process can be performed. An important thing to note here is that whether a point is declared implausible or non-implausible is not affected by these approximations. This correction is only so the iterative procedure can work and, although heuristic, it has worked on a large number of entities and experimental setups, finding non-implausible points. The process of finding a non-implausible point is essentially the same as Algorithm 9. The difference is that instead of history matching $\mathcal{F}'$, $\mathcal{F}'$ is passed through the transfer function. The variance of the

observation minus the adjusted expectation of the reified simulator, notated σ_{md}, is shown in Figure 6.8. A small correction is applied to the image before applying Algorithm 12. The summary of the scatter removal method developed in the last three chapters is given by Algorithm 22.

Algorithm 22 Scatter Removal Method: Online Calculations

Inputs: z, Direct-Beam-Emulator, $\mathcal{LU}$, $\hat{\beta}^{\Delta}$, σ_{md} , ImThesh

Outputs: $\hat{x_S}$, $\chi_s^\star$, $z - \mathbb{E}_{D^f, D^{\mathcal{F}}}[\mathcal{F}^\star(\hat{x_S}, x_E^\star, w^\star, M = 2)]$

1: Apply Algorithm 10 to create Prior-Specification-for-$x_S^\star$.

2: Using Prior-Specification-for-$x_S^\star$, $\mathcal{LU}$, Direct-Beam-Emulator, z, $\hat{\beta}^{\Delta}$, σ_{md} and ImThresh, apply Algorithm 9 to find non-implausible point $\hat{x_S}$ and set $\chi_s^\star$. Here the emulator is of the reified simulator, the correction in equation 6.5.3 is applied and the variance of the difference between the observations and the adjusted expectation of the reified simulator is σ_{md}.

3: Using $\hat{x_S}$, calculate the approximated adjusted expectation

$$\mathbb{E}_{\{D^f, D^{\mathcal{F}}\}}[\mathcal{F}^\star(\hat{x_S}, x_E^\star, w^\star, M = 2)],$$

up-sample and calculate scatter corrected image $z - \mathbb{E}_{\{D^f, D^{\mathcal{F}}\}}[\mathcal{F}^\star(\hat{x_S}, x_E^\star, w^\star, M = 2)]$.

return i) convergence plots and ii) implausibility and percentage error images.

6.6 Real World Results

6.6.1 Quantitative Phantom Study

The prior specification for the system values of the quantitative phantoms was made using a multivariate random field. The thickness was assumed to be in the range $[1, 12]$ and the alloy of PMMA and aluminium $[0.1, 1]$. The thickness field was specified as

$$\eta(i, j) = (1, i, j)\beta^\eta + \mathcal{W}^\eta(i, j)$$

where the covariance specification of $\mathcal{W}^\eta(i, j)$ was made using equation 2.2.8 with correlation length $l = [5, 5]$, $p = 2$ and $\mathbb{V}ar[(i, j)] = 0.1$. The regression parameters were fixed using GLS estimates, see key-formula 2.2.1. The alloy field was specified as a mixture of random fields, see equation 2.3.1, given by

$$\mu(i, j) = \sum_{\omega=1}^{2} \pi_{(\omega)}(i, j)(\mathcal{W}^\mu_{(\omega)}(i, j)).$$

A region in the image was specified manually as *a-priori* known. For the region with known material, $\mathbb{E}[\mathcal{W}^{\mu}_{(\omega)}(i,j)] = 1$ and $\mathbb{V}ar[\mathcal{W}^{\mu}_{(\omega)}(i,j)] = 0$. For the region with uncertain material, the correlation structure of the residual process was specified using equation 2.2.8 with $l = [0.1, 0.1]$ and $p = 2$. This specified that the thickness field is smooth and that the alloy field in the unknown region is essentially uncorrelated.

Specifying that a region of the image has a known material improves the efficiency of the method in that fewer evaluations of the emulator are required. There are problems when too few pixels are classified as known tissue but it is robust to not specifying areas which are tissue as *a-priori* tissue. It is not robust to specifying a region containing bone as *a-priori* tissue. This causes problems with adjusting beliefs for the thickness field as areas incorrectly specified as tissue have a spike in thickness to compensate for the incorrect specification. As tissue is more scattering than bone, this also induces an overestimate in the scatter estimates. The iterative process then fails to find a non-implausible solution. The current method is a manual specification of tissue regions using a user-interface. There has been work aiming to automate this process with a neural network classifier, but this is still some way off being able to replace user interaction effectively.

CDRAD Phantom

The first phantom is known as the CDRAD phantom. Figure 6.10 shows a diagram of the CDRAD phantom as placed on the detector. This phantom is designed to test image quality, particularly contrast and resolution. It is a 1cm thick PMMA cuboid with a number of lead wires inside which mark out the boundaries of a grid. Within each element of the grid is a small cylindrical hole. The size of the whole varies across and down the grid testing contrast-to-noise and resolution. The width and diameter of the holes vary between 0.3–8mm. To increase scatter-to-primary-ratio (see definition 3.5.1), an 8cm thick block of PMMA is added on top of the phantom. This makes the total thickness 9cm. Measurements were made with and without an ASG using the same experimental set up. Therefore, the dose is identical for all images. For the CDRAD phantom, the entire object was assumed to be PMMA and the thickness field was assumed to be quadratic.

Figure 6.11 shows the convergence statistics for the iterative process. The process finds a number of non-implausible points. The difference between iterations deceases to minor changes for the last two iterations. Figure 6.12 shows a non-implausible thickness field and the expected scatter given the non-implausible thickness field. Another good

feature of this process is illustrated by Figure 6.12. There is small spatial variation in the error in the thickness field. Even if there is no scatter, most methods have spatial variation caused by the spatially varying spectrum. However, allowing the transfer functions to vary spatially means that this spatial error is mostly calibrated out. Therefore, even if scatter is a small factor, this method is useful in calibrating out spatial affects.

The shape of the scatter is as expected as scattering behaviour is dominated by the 8cm block of PMMA placed on top of the phantom. Figure 6.13 compares the i) original ii) scatter corrected image using our algorithm and iii) image with an ASG. The scatter corrected image is superior as it is flatter, which is a better representation of the underlying object. The cylindrical holes and the surrounding regions are visible for all elements. The ASG on the other hand has a clear shape remaining from scatter. The holes are beginning to be obscured in the centre and at the edge of the image. The original is far worse than both ASG and scatter corrected images. This is also shown by the image histograms. The peaks for the scatter corrected version is sharper than the ASG and the original. In terms of image quality, the scatter correction algorithm out performs the ASG and the original. This implies that the images produced from this algorithm can be better quality than the current state-of-the-art used in industry.

These results have been visually verified by industry leaders comparing the images produced by our software and their own methods using the same phantom. This has been done by a number of companies using a number of detectors and a number of experimental set ups. There has also been verification using third party software which analyses the CDRAD phantom. Due to commercial considerations, of our industrial partners and their customers, we are not able to i) present specific results or ii) quote individuals or their companies.

Arm Phantom

The second phantom is referred to as the arm phantom. The geometry is shown in Figure 6.14 and it was designed to assess whether non-implausible points are useful for inferring entity composition. The arm phantom was placed on the detector with the flat side facing the source. The centre cylinder was filled with aluminium. Images were taken with and without the ASG at the same experimental settings. Figure 6.15 shows the convergence plots for the arm phantom. The implausibility drops quickly to a non-implausible point and then hardly changes for 5 iterations. The standardised errors jump between under and

over estimates until convergence. Figure 6.16 shows a non-implausible thickness and alloy field, and the standardised residuals for line profiles. The solution is within 10% percent of the known truth for the majority of the image and the standardised residuals show no evidence the correct answer is not within the non-implausible sets. The standardised residuals are small which indicates that the uncertainty may have been overestimated. Figure 6.17 shows the percentage error between a non-implausible point and the known truth for a profile running down the centre of the entity. The percentage error is in $[\pm 10]\%$ which shows the entity composition returned can be used to make inferences for aBMD. Figure 6.18 shows an image comparison of the region containing aluminium. For SRM, the most detail is visible and the majority of the shape from scatter has been removed. The ASG is again superior to the original as, for the same contrast stretch, the original has saturated towards the edges.

Arm Phantom 2

To continue the quantitative study, a set of 6 simpler phantoms were used. The simpler phantoms were chosen as it was easier to vary the alloy present [3]. The phantom design is given in Figure 4.6 top right. The phantom was made out of PMMA and aluminium. The total thickness of every entity in the set of phantoms was 5cm. The thickness of aluminium present in the centre of the phantom was $[0.5, 1, 1.5, 2, 2.5, 3]cm$. Figure 6.19 shows the percentage error between a non-implausible point and the known truth. The values in the plot are the mean of the absolute value of the percentage errors in the region on the detector, where the entity above contained aluminium. The far right plot shows a point estimate for the thickness of aluminium. If the measure was of bone thickness as opposed to aluminium, it can be transformed directly into aBMD.

6.6.2 Cadaver Study

When moving onto performing the analysis for cadavers, there is an additional source of discrepancy. This is due to the fact that the chemical composition of tissue and bone is more uncertain than that of PMMA and aluminium. Furthermore, whether the same chemical composition is appropriate for all patients is doubtful. The analysis is performed using G4 soft tissue and G4 cortical bone. We expect this to match real world soft tissue

[3] They were cheaper!

and bone less well than G4 PMMA and G4 aluminium. The assumption made, see equation 6.2.3, is that for all $x_S^\star$ the simulator performs equally well. For cadavers, further work is required to validate this assumption or to propagate the error induced through to our uncertainty analyses. However, as will be shown in this section, even with this additional discrepancy, image quality can be improved and the composition inference looks plausible in that it is correlated with our beliefs about the composition of anatomy examined. As the ground truth is not known for cadavers, the accuracy of composition inference is not assessed.

The prior specification for the system values of the cadaver study was made using a multivariate random field. The thickness was assumed to be in the range $[1, 20]$ and the alloy of G4 tissue and G4 cortical bone $[0.5, 1]$. The thickness field was specified as

$$\eta(i,j) = (1, i, j, ij, i^2, j^2)\beta^\eta + \mathcal{W}^\eta(i,j)$$

where the second order specification of $\mathcal{W}^\eta(i,j)$ was made using equation 2.2.8 with correlation length $l = [15, 15]$, $p = 2$ and variance multiplier 0.1. The regression parameters were fixed using GLS estimates, see key-formula 2.2.1. The alloy field was specified as a mixture of random fields given by

$$\mu(i,j) = \sum_{\omega=1}^{2} \pi(i,j)_{(\omega)}(\mathcal{W}^\mu_{(\omega)}(i,j)).$$

A region in the image was specified manually as *a-priori* tissue. For the known region, $\mathbb{E}[\mathcal{W}^\mu_{(\omega)}(i,j)] = 1$ and $\mathbb{V}ar[\mathcal{W}^\mu_{(\omega)}(i,j)] = 0$. For the region with uncertain material, the correlation was specified using a squared exponential with $l = [0.1, 0.1]$.

To test the algorithm on clinically relevant samples, images of a cadaver were used. The study took place at Newcastle Clinical Training Centre (NSTC) and images of a knee, NOF and lumbar spine were used to give a range of clinically relevant images. Given that for the cadavers the true entity composition is unknown, the purpose of the cadaver trial is to assess image quality. Here we will compare original, ASG and the scatter reduction method (SRM). This is performed using the contrast comparison method described at the beginning of this section. This image quality analysis enables all three images (original, ASG and SRM) to be compared for the same contrast window around the feature of interest.

Knee

Figure 6.20 shows a diagram of the region imaged of the cadaver's knee. Figure 6.21 shows the outputs of the algorithm for the knee. The expected thickness field is reasonable given manual measurements of the cadaver, although these measurements were not very precise. The alloy field shows appropriate correlation in that regions expected to contain more bone are a smaller alloy. Figure 6.22 shows a comparison of the three images. By this method, the image quality is best for SRM, ASG then original. The improvement is most clear on the lower portion of the knee where, for the same contrast window, the detail is lost for the ASG, as it is saturated in that region, and a similar effect can be seen on the edges. The improvement is minor as the proportion of intensity due to scattered photons is small. However, this shows that, for a knee, scatter can be removed without damaging image quality.

NOF

Figure 6.23 shows the region of the cadaver imaged. Figure 6.24 shows the outputs for the NOF . The thickness field is plausible with some underestimation in the centre. The alloy field looks reasonable but the bone regions may be underestimated. The shape change in alloy in the centre is also concerning. We believe this is due to the model of the thickness field being too simple for the structure of a NOF. Figure 6.25 shows the image comparison for a NOF. This time there is a more obvious improvement between the ASG and original which is why ASGs are often used for images of NOFs. For the same contrast window, there are large regions which are no longer visible, particularly a large area which has been saturated as white and black showing no detail of the bone structure. SRM is superior in that most of the image is visible and it has the best fine detail. This is clearest above and below the joint as detail remains after the colour has been saturated on the other images.

Lumbar Spine

Figure 6.26 shows the region of the body imaged. Figure 6.27 shows the outputs of the algorithm for the lumbar spine. The thickness is again plausible given measurements of the cadaver and the alloy field shows correlation with expected bone density. Figure 6.28 shows the image comparison for the lumbar spine. The lumbar spine has the highest proportion of scatter and, as such, shows the most obvious differences. The bone detail is difficult to make out in the original and the ASG is a large improvement with most of the

bone structure becoming visible. The protrusions from the spine are only just visible and the detail on some of them is lost. The SRM image has all protrusions clearly visible and overall better detail. This is further shown by the image histograms which show that most of the information of the SRM image is constrained in a smaller range. Figure 6.29 shows a contrast comparison between ASG and SRM with the mean of both images normalised to zero. The circles are to guide the reader to regions where the difference is most notable. The contrast has been set to show as much of the bone detail as possible. The ASG image has lost detail at both the top and bottom end resulting in features saturating as white or black.

Overall image quality is a hard thing to assess. The argument made here is that an image is considered better if for the same contrast window more information is available. In other words, less information is lost in the classification process when an image is produced. It is an open and clear way for experts to decide whether the overall image is improved and is less prone to bias than conventional methods like signal-to-noise ratio and contrast-to-noise ratio.

6.7 Conclusions

In this chapter, a method for transferring inferences from complex computer models to physical systems has been presented. The method assumes that one simulator is informative for multiple systems and measurements of calibration entities with known system values are available. The calibration data-set resolves uncertainty in i) the tuning parameters and ii) the difference between the tuned simulator and its reified counterpart. This gives a method for using a data-base created in simulation to make inferences about real-world physical phenomena and enables a solution to problem statement 1.1.4.

The method was applied to a number of entities on a number of different experimental equipment. The non-implausible points were within [10]% of the known truth for the system values for quantitative phantoms. The quantitative phantoms and cadaver images showed an improvement in image quality over that of an ASG. The methodology presented offers a route into the use of Bayesian inverse problem solving in X-ray imaging for scatter correction and entity inference. This has the potential to improve image quality and return diagnostic measures for clinical X-ray imaging with no engineering alterations to the equipment and in some scenarios a reduction in dose, see Appendix A.8 for discussion on dose considerations.

In Chapter 6, the advantages of our approach to predicting scatter − in that the error to the simulator was smaller − were presented. In Chapter 5, the advantages of the history matching approach − in that i) estimates for material and thickness (not just thickness) are returned and may be used to calculate diagnostic quantities and ii) an absolute (not relative) measure of whether an acceptable match to an X-ray image has been attained is used − were presented. In this chapter, the advantages of considering model discrepancy have been presented: the error between the simulator and the real world is reduced and its irreducible portion quantified. Therefore, the advantages of our approach to scatter correction are: i) the method can predict scatter more accurately, ii) whether the solution is *good enough* is well tested, ii) diagnostic measures are returned from a mono-spectral X-ray image and iii) the error between the simulator and the real world is smaller and its irreducible portion is quantified.

There were a number of avenues which we believe could be fruitful for future research. The first is a spatially dependent input spectrum. The solution we have given to this problem is to allow the transfer-function to vary spatially. The approach that is most commensurate with beliefs about the real world system would be to allow the spectrum entering the entity to vary spatially. This would involve making x_E into a field and changing the hierarchical-superposition emulator to take the field x_E as an input. The approach could still use pencil beams but there is likely to be complication caused by the increased dimensionality.

Another problem is that, although we assumed that bone and tissue is as well modelled as PMMA and aluminium, it probably is not. The extent to which G4 matches tissue and bone needs to be assessed. The extent to which the discrepancy varies patient to patient needs to be assessed. A method for resolving the uncertainty caused by this discrepancy and quantifying its irreducible portion is also required. A possible approach would be to use computed-tomography (CT) scans of a patient and compare the thickness and alloy reported, to inferences performed on projections. CT enable accurate measures of thickness so the system values for those cadavers would then be better known.

In this thesis, the experimental set up was held constant. However, performing reification training for every possible experimental set up is impractical. Therefore, a method for making inferences for a reified simulator with variable $x_S^\star$ would aid adoption in clinical practice. The structural tuning analysis could be used to assess whether changing experimental parameters can be tuned out.

CT scans are also affected by scatter so adapting the technique to make inferences using multiple projections could have a large impact on clinical practice. The inverse problem solving approach also has potential to solve classic problems in CT like beam hardening artefacts. The major barrier is the increase in dimensionality both of input and output space. A multi-resolution history matching approach is our suggested started point.

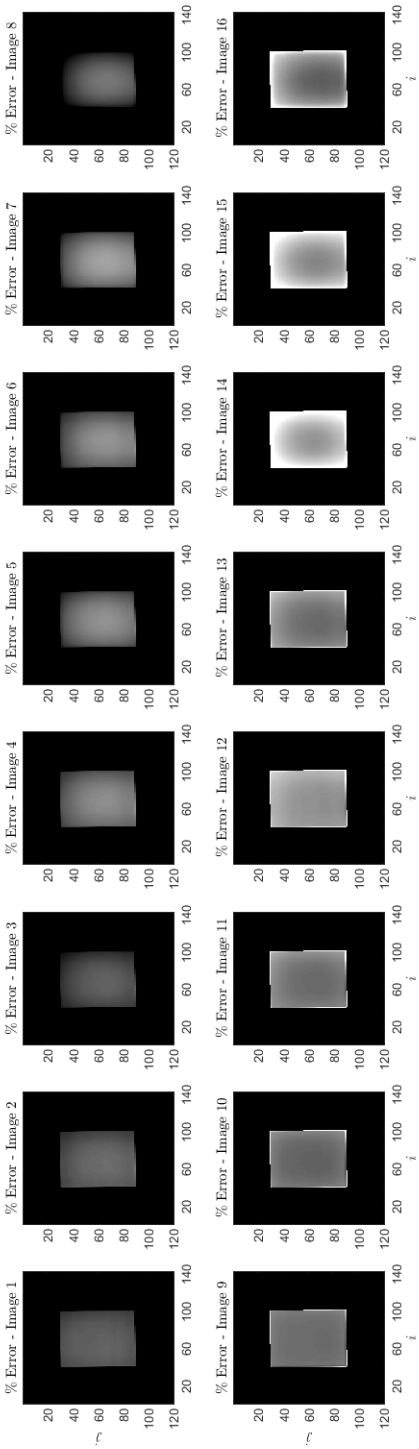

Figure 6.7: Percentage error for every entity in the calibration training-set. The colour-map is scaled to $[\pm 5]$.

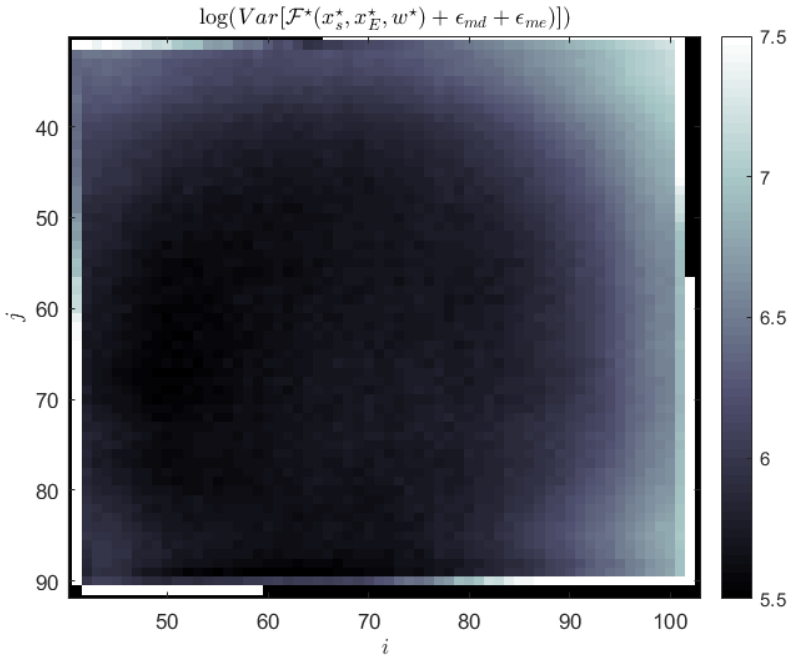

Figure 6.8: Image of the diagonal components of the sample covariance matrix of difference between i) the adjusted expectation of the reified emulator and ii) the calibration observations.

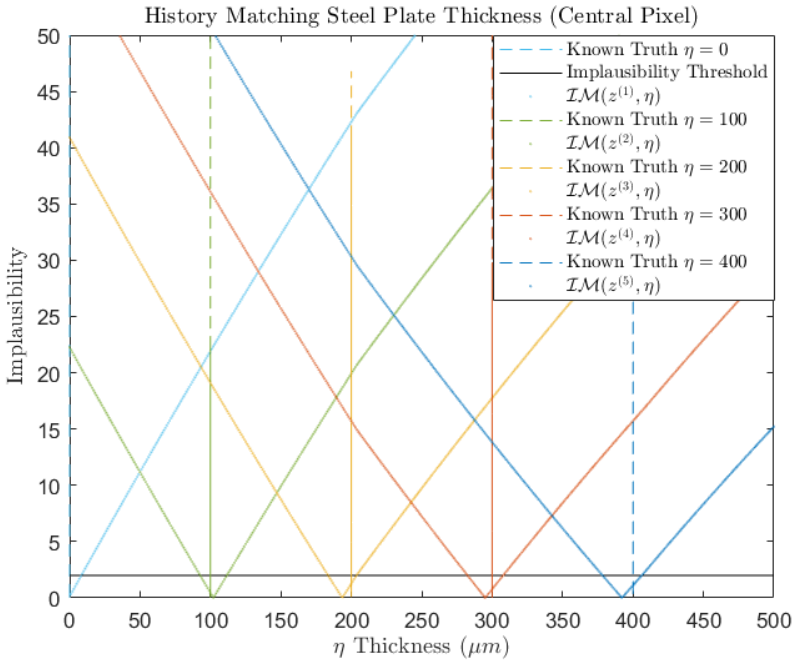

Figure 6.9: Figure showing the calculations of thickness non-implausible sets for five thicknesses for stainless steel $\{\eta^{\star(i)} = (i-1) * 100\}_{i=1:5}$.

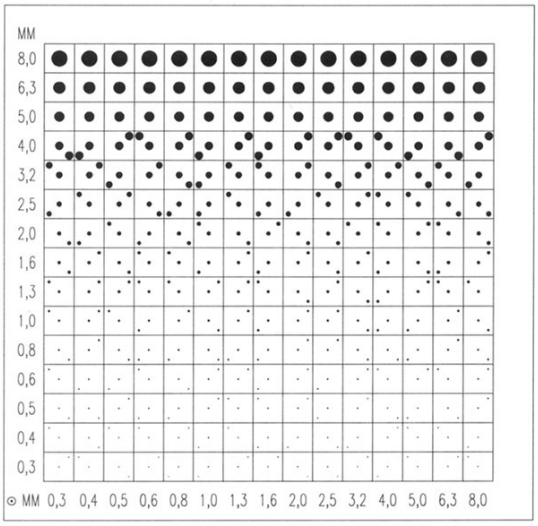

Figure 6.10: Diagram of the CDRAD phantom as seen by the detector.

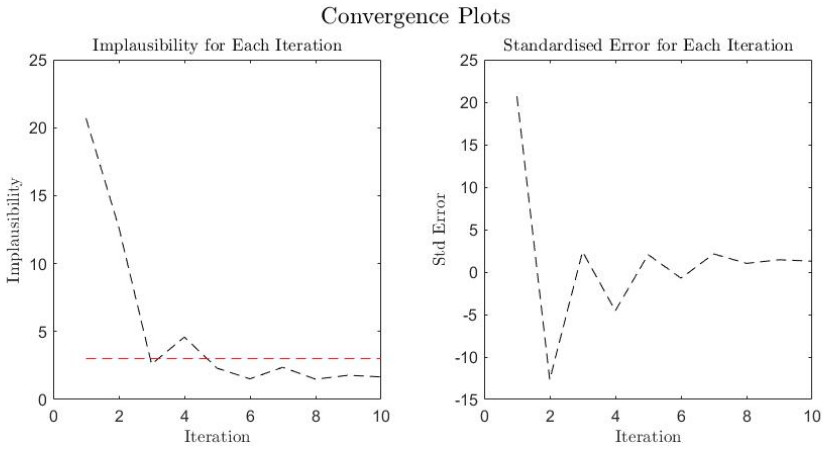

Figure 6.11: Left: implausibility measure for every iteration. Right: standardised residual for every iteration.

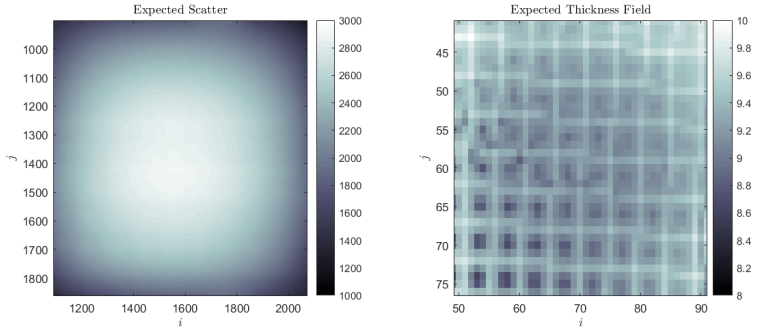

Figure 6.12: Left: non-implausible scatter removed for scatter correction. Right: non-implausible thickness field.

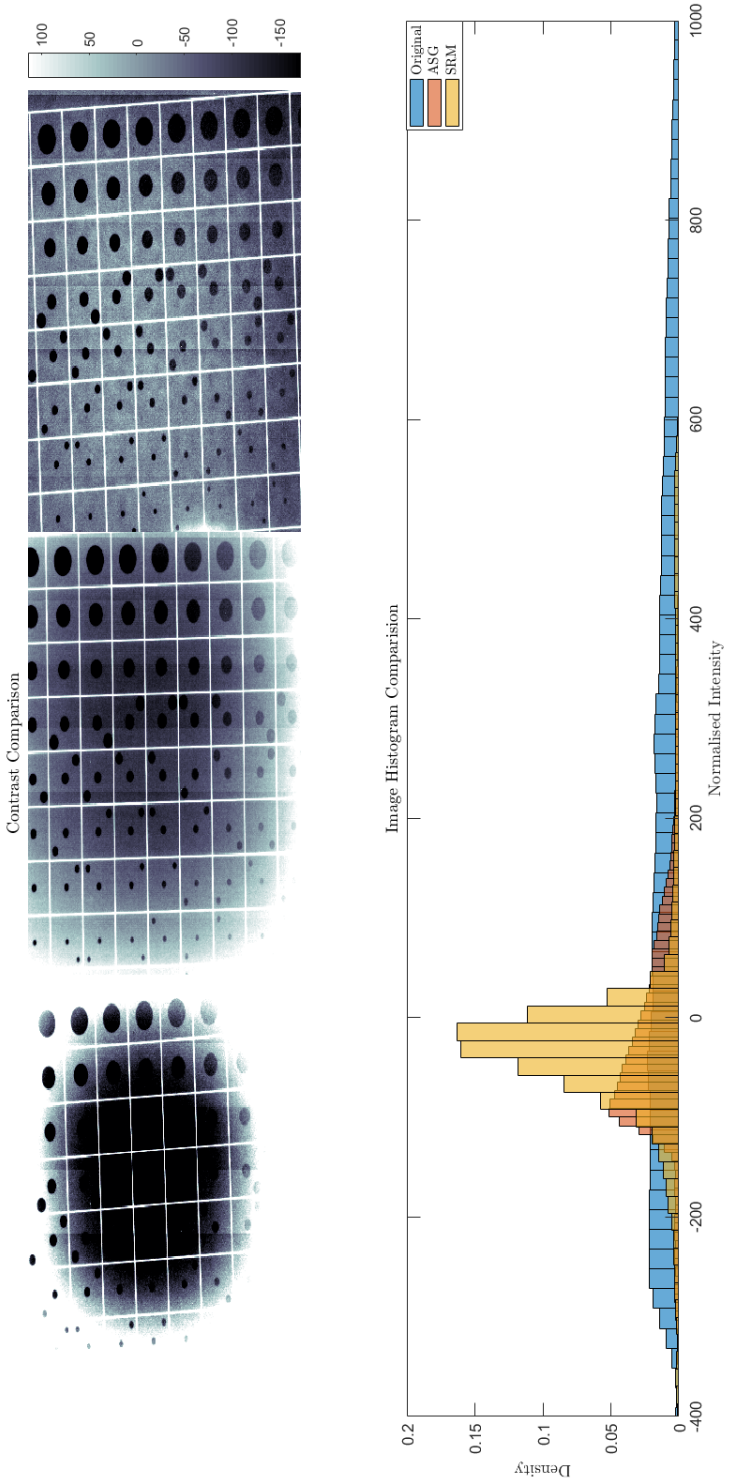

Figure 6.13: Left: image comparison of, from bottom to top, i) original, ii) ASG and ii)scatter removal method. Right: Histogram comparison of i) original, ii)ASG and iii)scatter removal method.

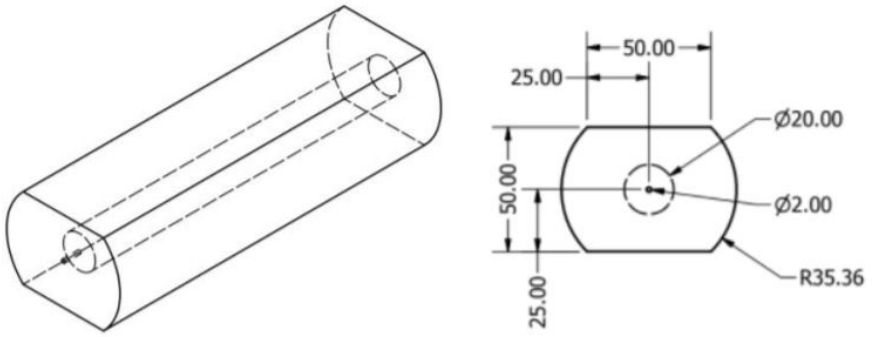

Figure 6.14: Diagram showing the geometry of the arm phantom.

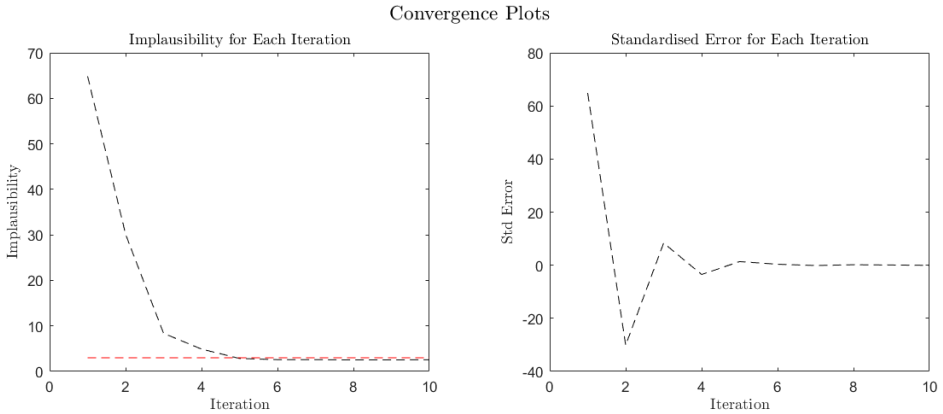

Figure 6.15: Left: implausibility for the arm phantom for every iteration. Right: standardised error for the arm phantom for every iteration.

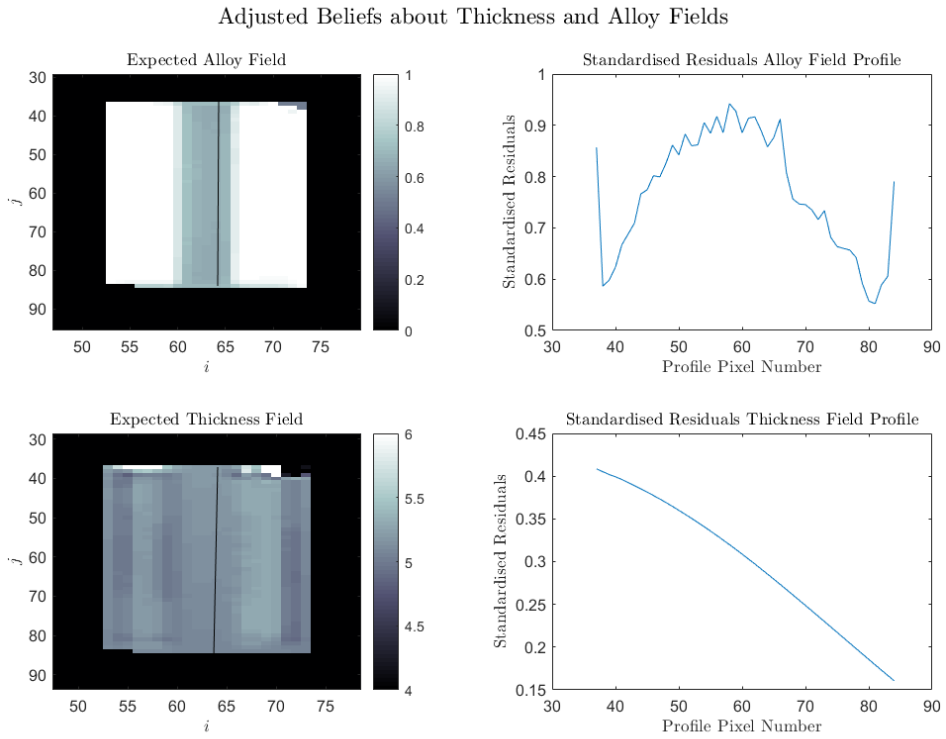

Figure 6.16: Top left: non-implausible point for the alloy field. Top right: standardised residuals for a profile of the alloy field. Bottom left: non-implausible point for thickness field. Bottom right: standardised residuals for a profile of the thickness field.

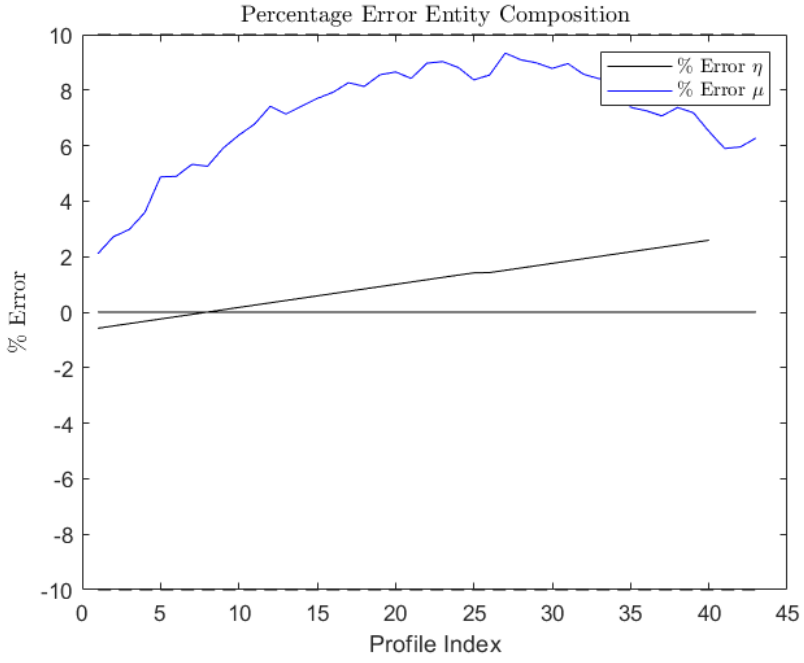

Figure 6.17: Percentage error between non-implausible point and known truth for profiles of the arm phantom.

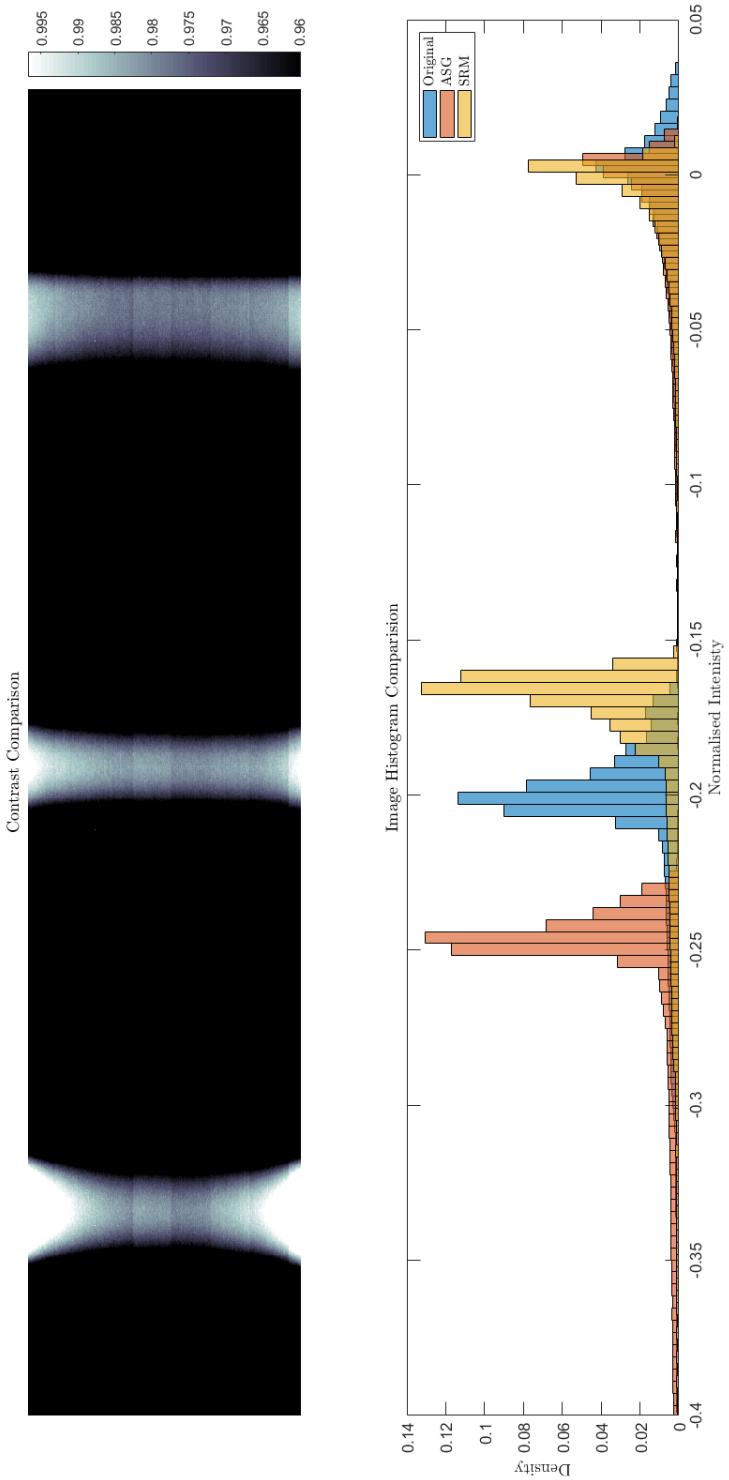

Figure 6.18: Left: image comparison of, from bottom to top, i)original, ii)ASG and iii)scatter removal method of the arm phantom. Right: histogram comparison of i) original, ii) ASG and iii) scatter removal method.

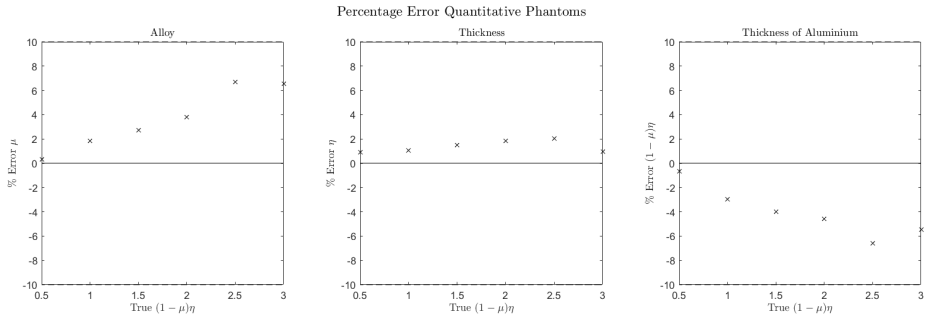

Figure 6.19: Left: average percentage error alloy. Centre: average percentage error thickness. Right: average percentage error aluminium thickness.

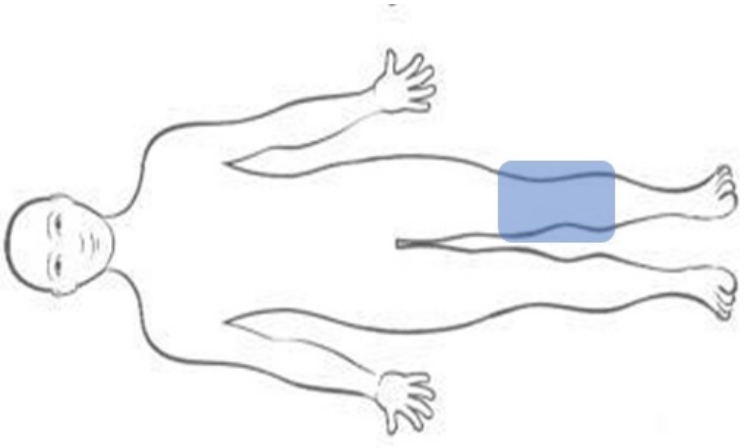

Figure 6.20: Diagram showing the region-of-interest for the knee image.

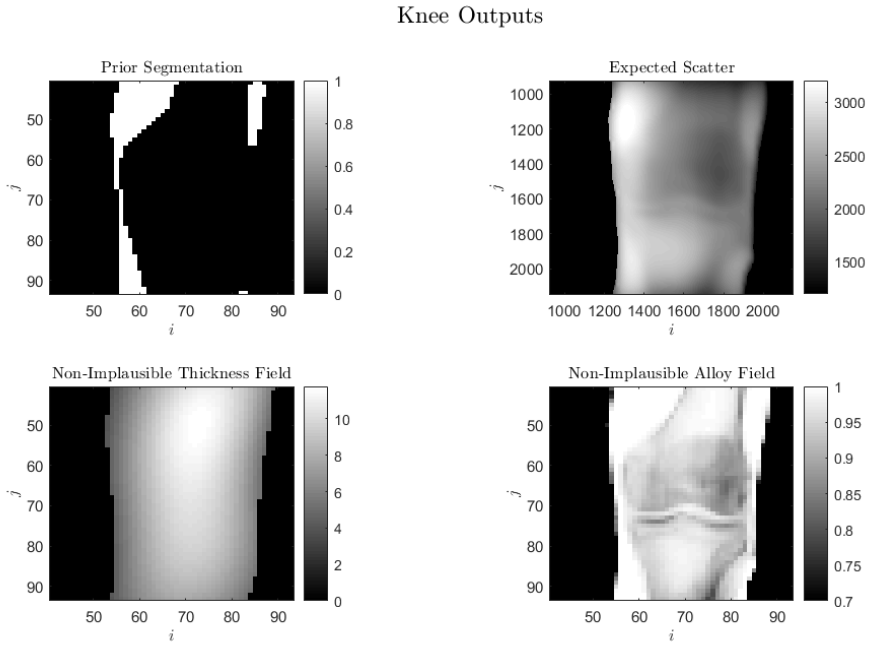

Figure 6.21: Top left: prior non-implausible tissue region for knee. Top right: expected up-sampled scatter. Bottom left: non-implausible thickness field. Bottom right: non-implausible alloy field.

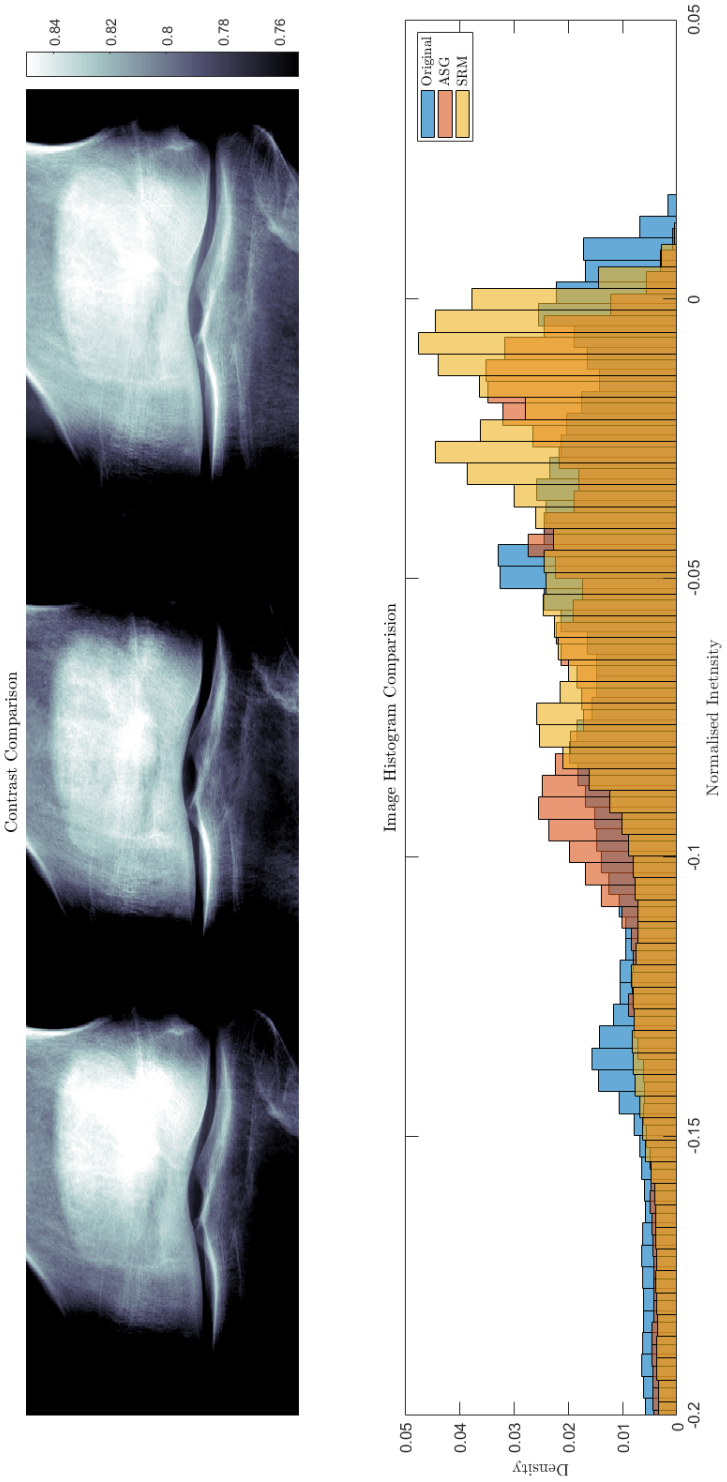

Figure 6.22: Left: image comparison of, from bottom to top, i) original, ii) ASG and iii) scatter removal method. Right: histogram comparison of i) original, ii) ASG and iii) scatter removal method.

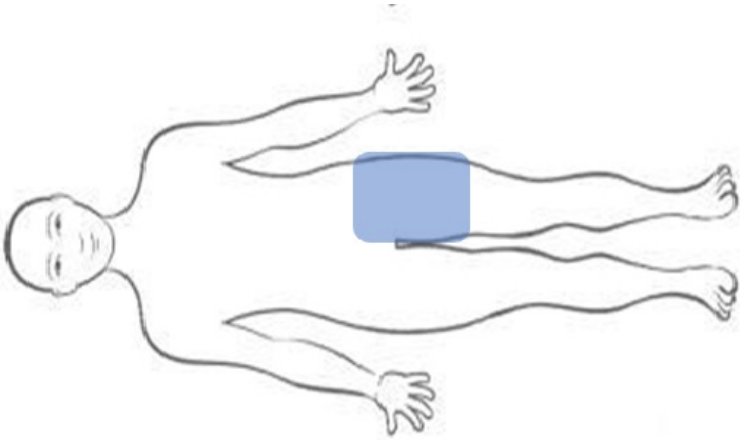

Figure 6.23: Diagram showing the region-of-interest for the NOF image.

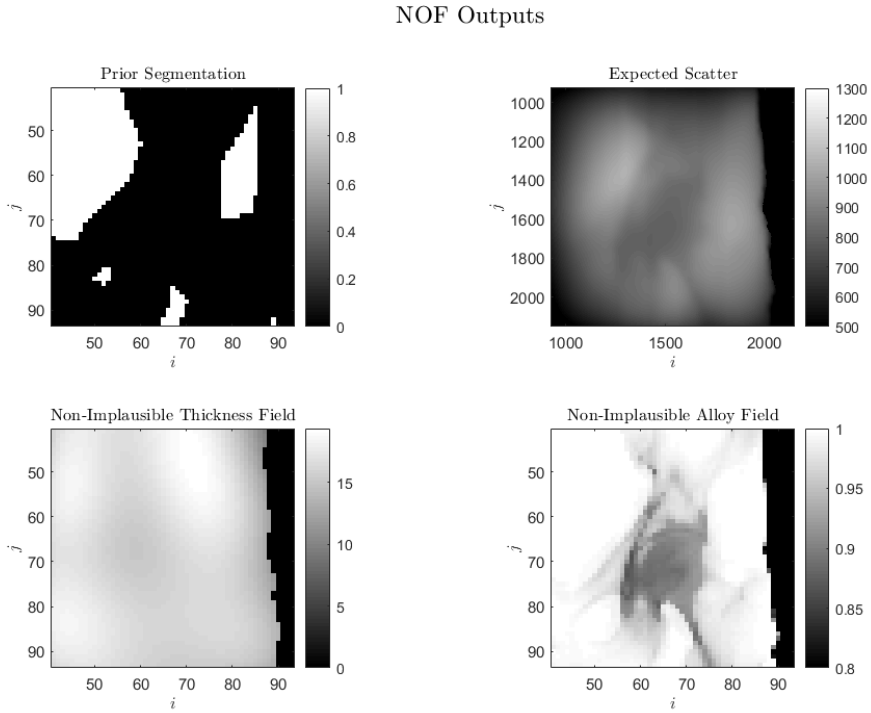

Figure 6.24: Top left: prior non-implausible tissue region for the NOF. Top right: expected up-sampled scatter. Bottom left: non-implausible thickness field. Bottom right: non-implausible alloy field.

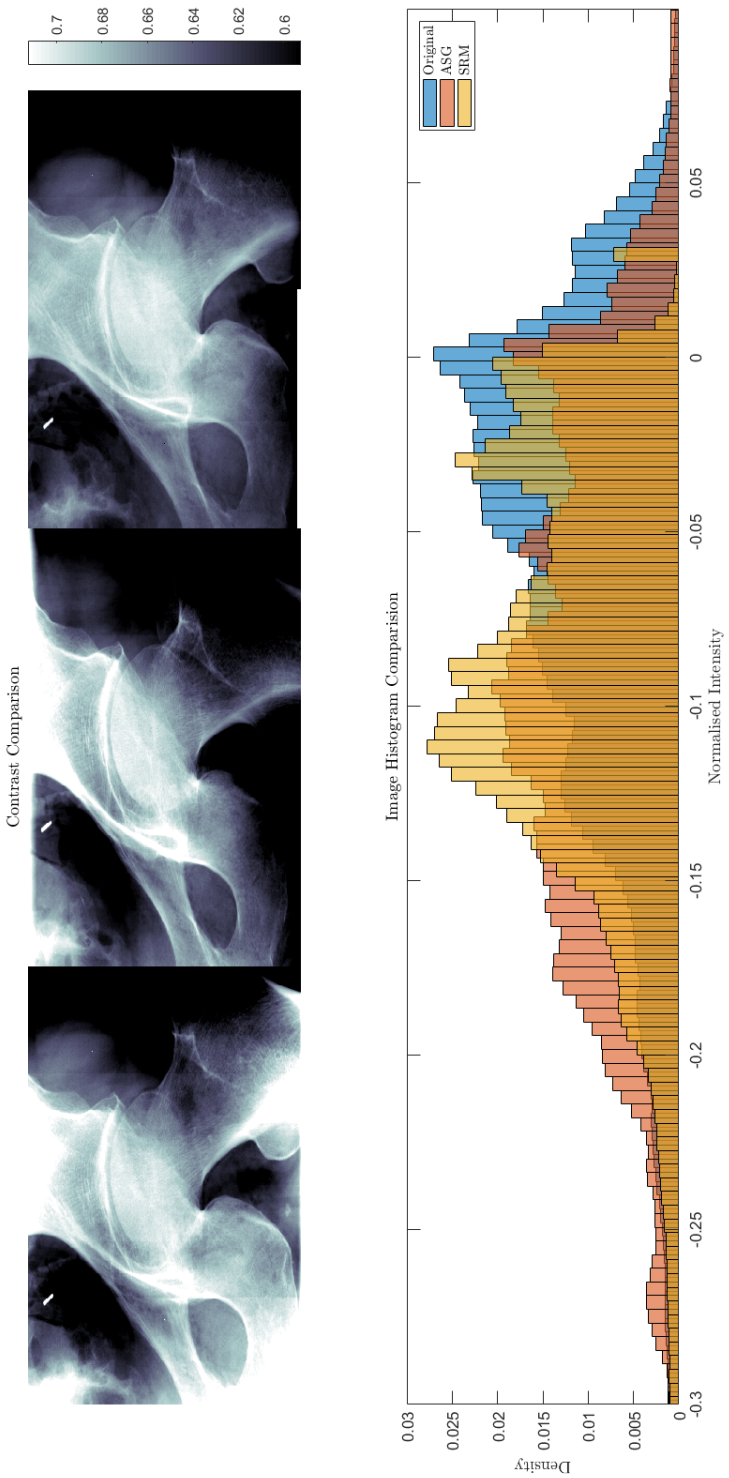

Figure 6.25: Left: image comparison of, from bottom to top, i) original, ii) ASG and iii) scatter removal method. Right: histogram comparison of i) original, ii) ASG and iii) scatter removal method.

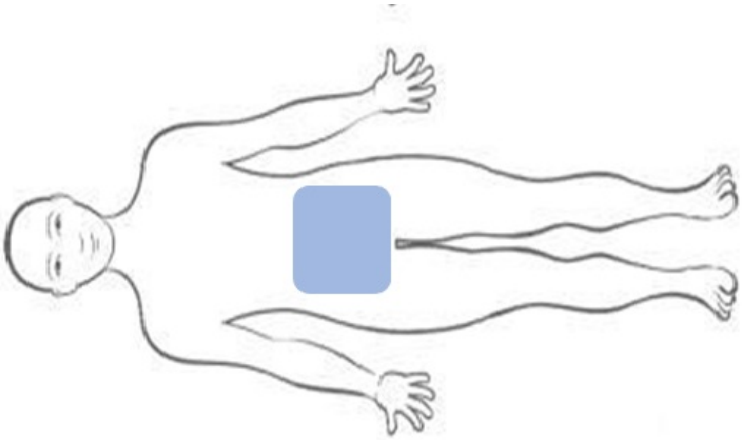

Figure 6.26: Diagram showing the region-of-interest for the lumbar spine image.

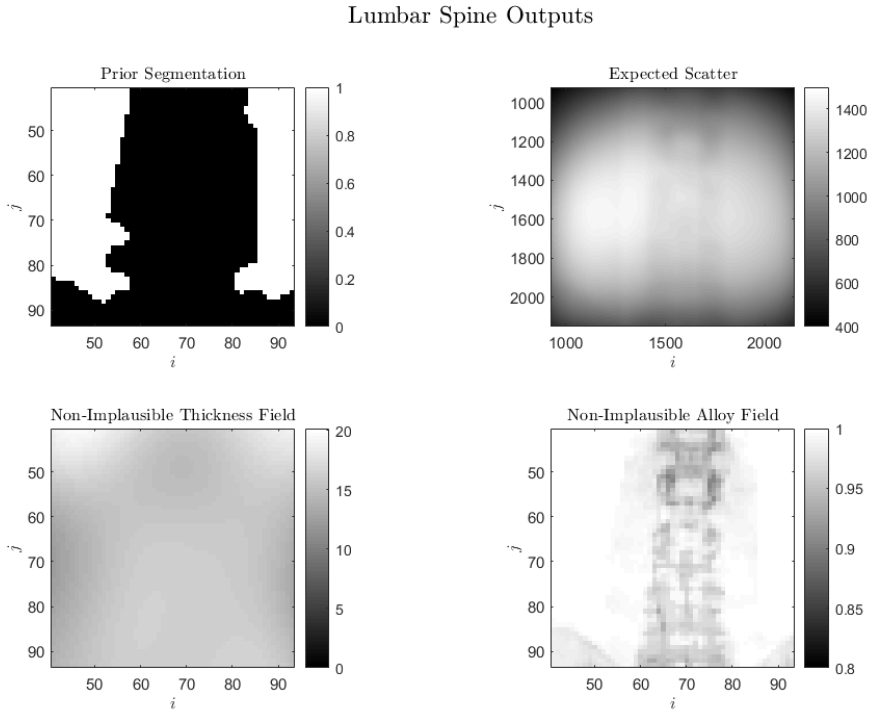

Figure 6.27: Top left: prior non-implausible tissue region for lumbar spine. Top right: expected up-sampled scatter. Bottom left: non-implausible thickness field. Bottom right: non implausible alloy field.

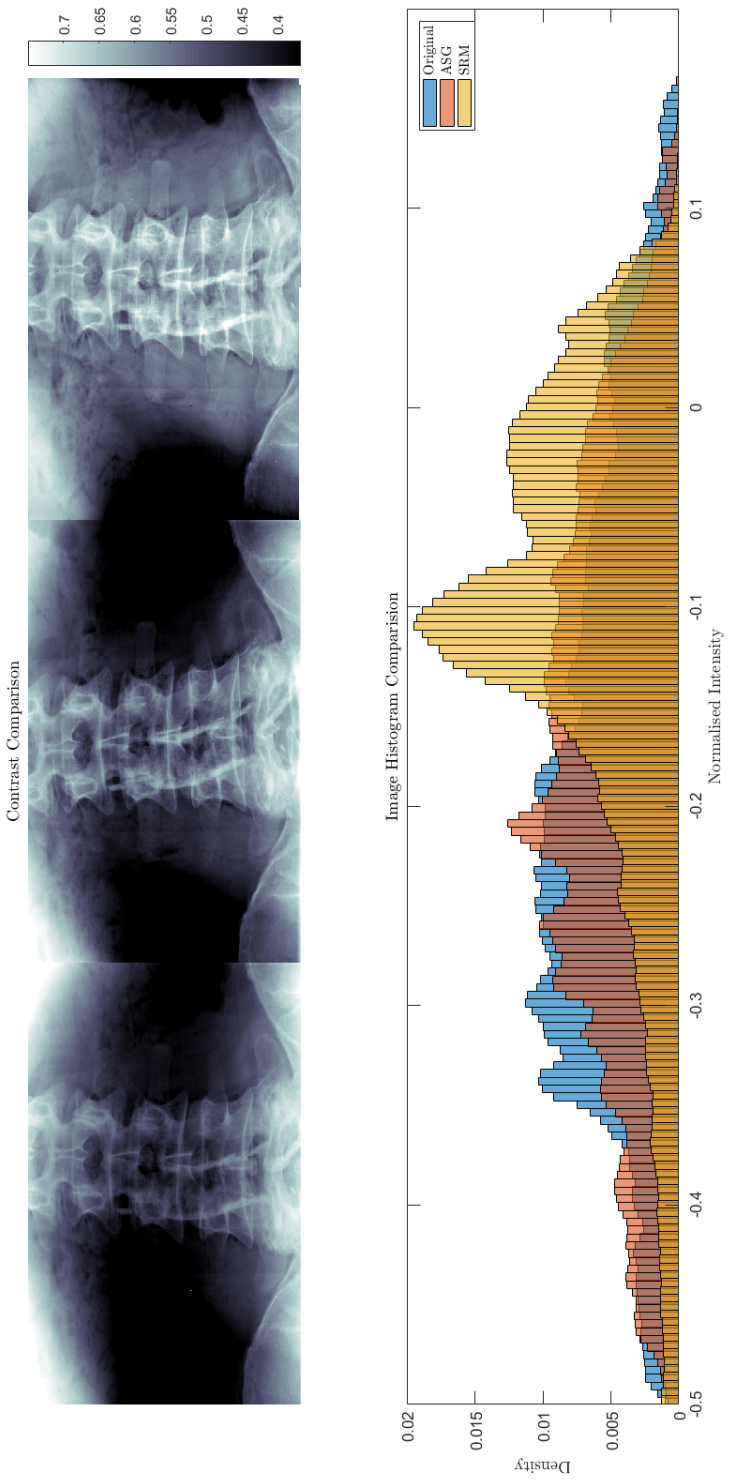

Figure 6.28: Left: image comparison of, from bottom to top, i) original, ii) ASG and iii) scatter removal method. Right: histogram comparison of i) original, ii) ASG and iii) scatter removal method.

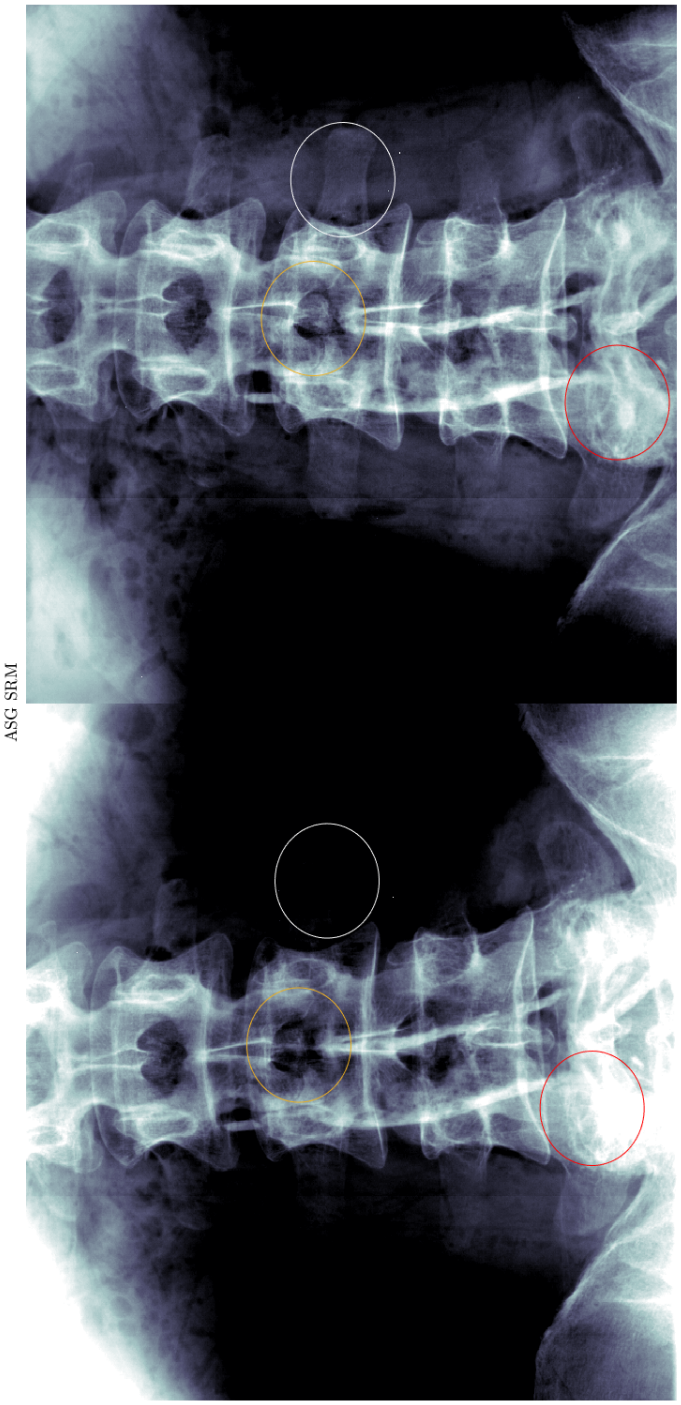

Figure 6.29: Bottom: image comparison with ASG. Top: image with scatter removal method applied.

Chapter 7

Conclusions

It is now time to move onto the conclusions of this thesis. What worked well, what was not successful and what future work would be required to facilitate adoption in industry.

7.1 Anomaly Detection

Chapter 2 presented the inferential paradigm used throughout this thesis, Bayes linear Methods. It enables prior judgements to be incorporated without a computationally challenging full Bayes analysis. In Chapter 3, the paradigm was illustrated by an anomaly detection methodology. Multivariate random fields enabled inferences for high dimensional random quantities. Second order exchangeability enabled inferences about the populations to which the random fields belong. Approximate kernel methods enabled inference with observation vectors containing many elements. Discrepancy measures enabled comparison between new observations and the non-anomalous population. Together this gives a tractable and efficient set of tools which may be used to solve anomaly detection problems.

The method developed is flexible but there are areas in which extensions to the methodology would aid its adoption in a wider range of applications.

1. **Classification.** Observations from both anomalous and non-anomalous populations may be available. For example, if the aim is to detect suspicious tissue in mammography, mammograms from both healthy and unhealthy populations may be available. Inferences could then be made for both populations and belief comparison performed to aid decision making. How inferences may be made given adjusted beliefs about two populations, including the relative population sizes, is an interesting question for future research into Bayes linear methods.

2. **Co-Exchangeable Multivariate Random Fields.** There may be additional meta-data available which changes beliefs about the multivariate random fields. For example, the age and weight of a patient may change beliefs about a mammogram of the patient. How this may be incorporated into the method is an interesting question and the extension would increase the range of problems which can be approached using this methodology.

3. **Uncertain Partitions.** The mixture of multivariate random fields specification assumed that the number of partitions is known. An interesting extension would allow uncertainty in the number of partitions. This would enable mixtures of multivariate random fields to be applied more generally to a wider range of observations.

The application of the anomaly detection methodology showed a proof of concept that plastic contaminants could be detected in chocolate bars using MAP equipped X-ray examination equipment. There were a number of features which made this problem challenging: the entities of interest were non-mean-differentiable, the observation vectors were very large and the signal from contamination was small. We know of no other approach which can detect plastic contaminants in chocolate using X-ray examination so the level of sensitivity achieved was seen as a significant result.

The algorithm has essentially two components: a spatial modelling step to identify suspicious regions and a spectral modelling step to decide if the suspicious regions are contaminated with plastic. The first component is possible on most X-ray examination equipment. The second is only possible on multi-spectral X-ray examination equipment. The spatial modelling component worked well and contaminants were consistently identified as spatial anomalies, even if they were barely visible in the images. The spectral test was more challenging as the chemical difference between plastic and chocolate is small. This part of the process consistently detected larger contaminants, 3-5mm. However, for smaller contaminates, the difference between contaminants and spatial anomalies was small and the decision was sensitive to small perturbations in parameters. The required specificity for a viable algorithm was quoted as 1 in 10000 chocolate bars wrongly identified as containing plastic. If the aim is to consistently detect 1mm contaminates with this level of accuracy, the signal induced by plastic would have to be increased.

7.2 Emulation and Scatter Prediction

Chapter 4 presented Bayesian emulation as a route to predicting the intensity in an X-ray image due to scattered photons. The method combined ideas from signal processing and Bayesian emulation to create an efficient model of a simulator where both inputs and outputs may be organised as multivariate random fields. The signal processing approach enables efficient modelling of the simulator but it requires strong assumptions. The Bayesian emulation formalism confronts these assumptions which enables an expert to think carefully about whether they are appropriate and how they may be relaxed. An extension which removes the usual signal processing assumption, namely piece-wise stationary and isotropic point spread functions, was proposed. The cost of relaxing the usual assumptions is that the computational benefits of the fast Fourier transform are lost. To enable efficient inference, look-up-tables were used. To include all available simulations in building a look-up-table, an emulation strategy was designed which enabled information from a large number of simulations to be incorporated in the adjustment. Together this enabled more accurate emulation of simulators, with random field inputs and outputs, in a reasonable time-frame.

The methods constituted a considerable step in the right direction, however there are areas where it could be improved.

1. **Uncertainty Estimation.** One part of this method that was not good was the uncertainty estimation. This was mainly a computational problem as the number of adjusted covariance calculations required was huge. Another cause was that the order of materials was assumed to be insignificant. The uncertainty induced by this assumption was assessed but not included in the inferences moving forward. Finally, the entities in G4 were relatively simple, therefore, when moving to more complex entities, additional uncertainty is likely to be incurred which should be assessed.

2. **Three Dimensional Model.** The entity geometry considered was essentially a two dimensional model. An extension would be to consider a three dimensional model with a material for every voxel. This would relax the assumption that the simulator output is invariant to the layer ordering of materials, enable the methods to be applied to computed-tomography, and facilitate a less heuristic ray tracing algorithm.

3. **Multi-Scale Emulation.** The inferences were performed with lower resolution

detectors than usually used in medical X-ray examination. This was sufficient as scatter is a low frequency signal so a simple up-sampling provided an improvement to image quality. However, lower resolution emulators will be more efficient and informative for higher resolution emulators. Therefore, an interesting extension would be to use multi-scale emulation methodology to improve the emulators efficiency and accuracy.

Emulating GEANT4 was challenging for a number of reasons: i) the simulator was expensive, ii) the input was high dimensional and iii) the simulator was not flexible enough to model entities of the required complexity. The method presented proposed a hypothetical improved simulator which was able to encode more flexible entities. A statistical model of this emulator was built using a linear combination of pencil beam simulations directed at carefully designed entities. The adjusted expectation of a simulator with $2(88 \times 105)$ inputs and (88×105) outputs could be calculated using i) (88×105) evaluations of an emulator with 2 inputs and scalar output and ii) $2(88 \times 105)^2$ evaluations of emulators with 5 inputs and scalar output. These calculations could be performed efficiently with look-up-tables which enabled inference in a time-frame which would not inhibit clinical practice. The emulator built had a lower error to GEANT4 than the current state-of-the-art approaches. The emulator was sufficiently accurate that the subsequent goals of the study − to reduce the effect of scattered photons and make inferences about an entity's composition − could be achieved.

7.3 History Matching with Random Field Priors and Entity Inference

Chapter 5 presented the history matching approach to Bayesian inverse problem solving. History matching offers a route to resolving uncertainty in the system values using computationally inexpensive calculations. How the outputs of the history match may then be used to make inferences for the quantities of interest was considered and the wider role of history matching in subjective Bayesian analysis discussed. An extension to history matching was developed which incorporated beliefs, specified as mixtures of multivariate random fields, that reduced the effective dimension of the history match. This enabled non-implausible sets to be calculated even when the simulator had a large input dimension. Given the complexity of the samples and the speed of the emulator, an alternative

approach to finding a non-implausible point was required. This combined history matching and iterative techniques from signal processing to find a non-implausible point with fewer evaluations of the emulator. This enabled inference to be performed in a reasonable time-frame even for complex entities and expensive emulators.

There are a number of directions for future research that this work suggest could prove fruitful.

1. **Iterative History Matching as a Computational Tool.** Iterative history matching can find a non-implausible point with far fewer evaluations of the emulator. However, the exhaustive search has much more information. Therefore, an interesting extension would be to use an iterative approach to identify an interesting region of the input space and then perform an exhaustive search in the surrounding area.

2. **Multi-Scale History Matching.** It has already been stated that lower resolution emulators are informative for higher resolution ones. Similarly, lower resolution posterior non-implausible sets of multivariate random fields would be informative for higher resolution ones. Therefore, history matching at lower resolution could be used to build a more informative prior non-implausible set at a slightly higher resolution and iterating until the required resolution is reached. This means that the space to be searched is smaller and that information from lower resolution emulators may be passed to higher resolution ones, increasing the efficiency and/or accuracy.

3. **Populations of Non-Implausible Sets.** There may be many observations of similar entities available. Therefore, it may be appropriate to consider the non-implausible sets for each of these entities as samples from a population. If inferences about the population can be made, for example by specifying beliefs about the system values as a second order exchangeable set of multivariate random fields, this could be used to construct a more informative prior non-implausible set, reducing the space to be searched.

The history matching methodology was applied in simulation to the task of inferring the entity composition of a patient and removing the effect of scattered photons. An exhaustive search and an iterative method was performed. The exhaustive search results showed that the problem is not in the Bayesian sense ill-posed and that significant uncertainty in the entity composition could be resolved. The iterative results showed that

a non-implausible point could be found in a useful time-frame. The error between the non-implausible point and known truth was small and the image improvement clear. This showed that, given G4 is reality, Bayesian inverse problem solving methods provide a route into a previously intractable problem, namely making inferences for an entity's composition and a composition dependent scatter reduction method. This shows potential that image quality can be improved and extra information, like aBMD, can be made available without the need for multi-spectral X-ray examination.

7.4 Transferring Inferences to the Real World

Chapter 6 presented a method of transferring inferences from complex computer models to physical systems. A method for using a single simulator as informative for multiple physical systems was presented. The method considers tuning to a reified simulator where system values are allowed to drift in order to resolve model discrepancy. The problem with using subjective Bayesian analysis to make inferences for tuning parameters is that the quantities involved − about which an expert is asked to specify beliefs − are not operationally defined. A method termed structural tuning was devised as a systematic approach to assess the discrepancies involved and how the tuning parameters would change in order to take them into account. Structural tuning was designed to provide an *a-priori* specification for the quantities involved and a specification of the expected posterior beliefs. Then the modelling assumptions may be assessed, beliefs specified and comparisons made between expected and revealed posterior quantities.

Model discrepancy is a challenging subject and there were a number of areas in which future research could be profitable.

1. **Entity Dependent Discrepancy.** The discrepancies considered were assumed to be associated with the equipment dependent inputs. This essentially assumed that the simulator models all entities as well as it models the calibration training-set. A interesting extension would be to quantify the uncertain pertaining to the entity dependent inputs, include the statements in adjusted beliefs about the system behaviour and develop methods to resolve some of the uncertainty.

2. **Multiple Experimental Equipment.** The examples considered were constrained so that all measurements were made on the same experimental equipment and the same experimental settings. Therefore, an interesting extension would be to consider

how multiple measurements on multiple systems and multiple experimental settings could be used to make inferences for tuning inputs and transfer functions for all possible equipment and experimental settings.

The method developed was applied to create a solution to the transferable database problem in order to make inferences for the composition of an entity and reduce the effect of scattered photons in an X-ray image. A number of the scatter removal methods reviewed used simulators to make inferences for physical systems. None of them considered the discrepancy between their simulators and real world, so even accepting its existence is an improvement. The approach specified the key assumption that was required for a viable method, namely that the difference between simulator and reified counterpart is independent of the entity dependent system values given the output of the simulator evaluated at the tuning input. The method devised an implausibility measure to test this assumption. It was designed so that only points for which the direct tuning assumption, and match to observed data, is acceptable are in the posterior non-implausible set. The transfer functions were allowed to vary spatially to resolve some of the discrepancy caused by under-resolved experimental parameters. This enabled the percentage error between the adjusted expectation of the reified simulator and the observations to be reduced to under 5%. This gives two significant improvements over current approaches, the error is smaller and the magnitude of the error is quantified. This enables non-implausible entity dependent system values to be identified. This is a major advantage of this approach, for example, if a maximum likelihood estimator was returned, there would be no guarantee that it represented an acceptable match to observed behaviour given the current state of uncertainty.

The work performed illuminated further sources of discrepancy in GEANT4 which may reduce the difference between simulator and reified counterpart.

1. **Higher Resolution Equipment Dependent Parameters.** The spectrum incident on the entity was considered spatially invariant. In fact, it is expected to be spatially variant. Therefore, including x'_E in the hierarchical-superposition emulator would be a more appropriate representation of the system. The dimension of the tuning inputs would be significantly increased so history matching with random field priors might be a good place to start.

2. **Entity Dependent Uncertainty in System Behaviour.** The variance of the

system behaviour was considered independent of x_S. However, the variance should depend on x_S as, for example, simulator uncertainty is a function of x_S. This extension could enable more uncertainty to be resolved and entity dependent model discrepancy to be incorporated.

Chapter 6 also presented the full solution to the entity composition inference and scattering problems. This was a challenging problem as i) the simulator of the system was high dimensional and insufficient to model entities of arbitrary complexity, ii) the inverse problem was challenging to solve as the dimension of the system values was large and iii) there is discrepancy between the simulator and the system it purports to represent. The literature review revealed how practitioners have attempted to deal with these issues and what features would have to be improved to create a more viable method. From the review we identified three areas in which the current approaches were limited: i) the simulators used had strong assumptions, ii) the entity's composition was not directly considered and iii) the discrepancy between the simulator and the system was ignored. To improve on the current approaches: i) the hierarchical-superposition emulator was used to improve the model of the system ii) history matching with random field priors was used to make inferences about an entity's composition and iii) tuning was used to resolve the discrepancy between the simulator and reality.

The aim of the study was to produce an algorithm which could provide equivalent image quality to an ASG, and return inferences about an entity's composition which would be useful to clinicians, at a lower dose to the patient. If these aims could be achieved, the technology developed was predicted to have a high impact on clinical practice.

The image quality was consistently better than an ASG. This surpassed expectations and means that image quality improvements are available to clinicians. The method has been tested on around 20 entities on around 6 different systems, and consistently performs better than an ASG. This has been verified by third parties and engaged industry leading X-ray equipment manufacturers.

The inferences about entity composition resolve a large amount of uncertainty in both thickness and material of the entity. Given that this problem was considered by experts ill-posed and therefore unsolvable, this is considered a considerable achievement. The non-implausible points returned are with [$\pm 10\%$] of the system values which shows potential for these methods to provide clinically useful diagnostic measures. However, compared to the error reported on specialised multi-spectral X-ray examination equipment, it is a large

error. Furthermore, a non-implausible point is only part of a full inference and is not an expert's adjusted belief. As the simulator, spatial model and history match improve, it is expected that this error will reduce. However, whether it will reduce to the same level as specialised multi-spectral equipment, remains to be seen.

Even if the uncertainty resolution cannot be increased to the extent that it is equivalent to other equipment, the inferences returned can still be useful. This technology requires no engineering alterations to the equipment, therefore it is available on any piece of X-ray examination equipment. Instead of replacing the referral procedure, it can act to aid decision making of clinicians. For example, the entity inferences returned could be used to identify symptomless patients which may benefit from referral for a specialised scan. Furthermore, given that this information is returned from a routine scan, a lot more information about the symptomless population will be available. Overall, the inferences about the entity's composition are impressive in that they provide new capability on mono-spectral X-ray examination equipment; however, they are not yet accurate enough to replace multi-spectral examination equipment.

The dose argument is a hard one to make, see Appendix A.8. For some entities there is a clear dose benefit but for others there is an arguable dose deficit. The argument is complicated by the fact that the quality of the images compared is not equal. Given that the algorithm presented is better at removing scatter than an ASG, is it appropriate to consider the ASG as better because it has lower local noise? If that metric is accepted, an image with no scatter removal will always be superior. The counter argument is that, if local noise is lower, more information is in the image which can be viewed by stretching the contrast.

Another thought process is that, as uncertainty has been resolved in the entity's composition, it should be possible to use this to improve image quality. A simple thought experiment to support this argument is as follows. Consider a pencil beam simulation of a cuboid entity of constant thickness and material. If no scatter is allowed to hit the detector, the only pixel which will be non-zero is the central pixel. If scatter is allowed to hit the detector, many pixels will have a non-zero value. As both cases have the central pixel, it is clear that there is more information in the scattering case. Therefore, it might be possible to resolve more uncertainty in the entity composition and therefore provide improved image quality. The barriers to a method providing these improvements are dimensionality, computational power and model discrepancy. However, there is an

argument that uncertainty in the direct component could be reduced below measurement error by considering the images which contain scatter.

In this thesis, a number of methodologies have been developed and two algorithms proposed which show novel capability in a saturated field. Bayesian statistical approaches give new benefits to the X-ray industry. The algorithms are being developed into software that, we hope, will enable these benefits to be available to industry in the next couple of years. The methods open new areas of research into using Bayesian inference for making inferences from X-ray images and further research could enable further benefits in the near future.

List of Symbols

This list defines the notation that will is used within the body of this thesis.

$\mathbb{A}_z[x]$ Adjusted version of x by z, page 12

$h(\cdot)$ A vector of basis functions, page 45

A An indicator function describing whether an observation belongs to a class or not, page 45

ρ Control point for a mixture of multivariate random fields, page 28

$\text{Corr}(x, y)$ Linear correlation between x and y, page 35

$\mathcal{C}$ Correction function in an iterative algorithm, page 157

$C(\cdot)$ A correlation function, page 19

l A correlation length, page 20

$\mathbb{C}ov[x]$ Covariance of random vector x, page 10

$\mathbb{C}ov_z[x, y]$ Covariance of x, y adjusted by z, page 11

y_{AV} Inferred second order quantities of diagnostic quantities, page 284

$\mathbb{D}(\chi)$ Number of elements in array χ, page 144

$\delta_{(k,k')}$ $\delta_{(k,k')} = 1$ if $k = k'$ and 0 otherwise, page 32

$Dis(x)$ The discrepancy between revealed and expected value of a random quantity x, page 13

Std_T The transformed standardised errors between the expectation of a random vector and its value when revealed, page 14

$D^{\mathcal{F}}$ Set of evaluations of a simulator, page 87

$D^{\bar{\mathcal{F}}}$ Set of sample mean and variances of repetitions of evaluations of a stochastic emulator, page 292

E Energy, page 62

β^{EC} Continuous energy parameters, page 66

x_E Tuning (equipment) inputs, page 194

x_E^f Experimental parameters, page 15

$\mathbb{E}[x]$ Expectation of a random vector x, page 9

$\mathbb{E}_y[x]$ Adjusted expectation of x, adjusted by y, page 11

t Location on a random field in normed-vector-space $\mathcal{T}$, page 15

t^c Central pixel on the detector, page 108

I_q Identity matrix of size q, page 14

Y^{Dr} The direct component of an X-ray image, page 99

Y^{Sc} The scatter component of an X-ray image, page 99

$\mathcal{IM}$ An implausibility metric, page 139

$\oplus$ A symbol to denote the sum of *a-priori* uncorrelated random variables, page 16

$\mathbb{I}_a$ Indicator function for event a occurring, page 10

I Total energy arriving at the detector from photons passing through an entity, page 62

I_0 Total energy arriving at the detector, page 62

K Approximate Bayes linear adjustment kernel, page 40

$\mathcal{L}(\cdot)$ A link function, page 45

μ Attenuation coefficient or material label, page 62

μ^{MAP} Attenuation coefficient of the material of a Multi Absorption Plate, page 64

$\mathcal{SM}(X)$ Sample mean of X, page 203

$\mathcal{SV}(X)$ Sample variance of X, page 203

$\mathcal{M}(\cdot)$ Mean function: A multivariate random field describing beliefs of global variation of a multivariate random field, page 19

$\mathcal{M}_M$ Random field describing global variation of the population mean of a stochastic simulator, page 292

$\mathcal{M}_V$ Random field describing global variation of the population variance of a stochastic simulator, page 292

ϵ_{me} A random quantity describing measurement error, page 16

ϵ_{me}^G A random quantity describing the global variation of measurement error, page 17

ϵ_{me}^L A random quantity describing the local variation of measurement error, page 17

$\mathcal{N}(\mu, \sigma)$ A Gaussian distribution with mean μ and variance σ, page 31

$\tilde{r}$ Number of control points for a mixture of multivariate random fields, page 28

$N(E)$ Number of photons of energy E, page 63

N Number of observations, page 15

f Experimental examination system used to make an observation, page 15

z An observation of system behaviour, page 45

D^f A set of observations of systems behaviours, page 15

$\otimes$ Kronecker product of two matrices, page 20

q Number of fields in a multivariate random field, page 15

x Inputs to a simulator, page 83

$x^\star$ System values, page 135

$M(\cdot)$ Population mean of a random quantity, page 17

$M^G(\cdot)$ A random quantity describing global variation of the population mean of a random quantity, page 17

$M^L(\cdot)$ A random quantity describing local variation of the population mean of a random quantity, page 17

$\mathbb{P}(a)$ Probability of event a occurring, page 10

β Regression parameters, page 45

$R(\cdot)$ The residual between a random quantity and its population mean, page 17

$R^G(\cdot)$ Random quantity describing the global variation of a residual between a random quantity and its population mean, page 18

$R^L(\cdot)$ Random quantity describing the local variation of a residual between a random quantity and its population mean, page 18

$\mathcal{W}(\cdot)$ Residual process: A multivariate random field describing beliefs of local variation of a multivariate random field, page 19

$\mathcal{W}_M$ Random field describing local variation of the population mean of a stochastic simulator, page 292

$\mathcal{W}_V$ Random field describing local variation of the population variance of a stochastic simulator, page 292

$\mathbb{RV}ar_z[x]$ variance of x resolved by z, page 12

$\perp\!\!\!\perp$ Symbol to notate that two random variables are independent, page 94

S An entity, page 15

x_{ss} Parameters for the sampling distribution used in the inference, page 284

$\mathcal{S}$ A set of entities, page 15

y_{FOS} Second order belief specification from an automated specification algorithm, page 283

T A set of locations on a random field, page 15

$\mathcal{F}$ A simulator or complex computer model, page 83

$\mathcal{F}^r$ Required set of simulator evaluations to be able to describe the system behaviour, page 85

$\mathcal{SN}(a, b, c)$ A skew-normal distribution with mean a variance b and skewness c, page 284

x_{ps} Parameters for specification of second order beliefs, page 283

Std The standardised error between revealed and expected value of a random quantity, page 13

$\bar{x}^{\star}$ Summary statistic of a non-implausible set, page 144

y System behaviour, page 15

η Length through an object a photon must traverses to reach the detector, page 62

$\mathcal{T}$ A normed-vector-space, page 15

$\mathcal{T}$ Transmission of X-rays, page 62

$\mathcal{U}(a, b)$ A uniform distribution on the interval $[a, b]$, page 284

$\mathbb{V}ar[c]$ Variance of a random vector x, page 10

$\mathbb{V}ar_y[x]$ Adjusted variance of z, adjusted by y, page 11

α A weight in a weighted sum, page 145

W Weight matrix for generalised least squares, page 289

u A vector which parameterises the remaining dependence on the in-active inputs, page 97

$x_{(:)}$ Vectorisation of matrix x, page 279

B B-spline basis function, page 281

CH The convex hull of a set, page 203

P Partitioning function, page 28

diag(x) A function which returns a square matrix with diagonal elements x, page 26

D A function describing the discretisation process, page 85

SS A function which sub samples an image, page 163

F* A reified simulator, page 191

aBMD areal Bone Mineral Density, page 100

ASG anti-scatter-grid, page 100

DR Digital Radiography, page 64

G4 GEANT4, page 103

GLS Generalised least squares, page 25

GPU Graphics processing unit, page 90

MAP Multi Absorption Plate, page 64

mAs Milliamp seconds, page 210

NOF Neck of Femur, page 232

NSTC Newcastle Clinical Training Centre, page 231

SNR Signal-to-noise ratio, page 299

SOE Second Order Exchangeability, page 48

SOERF Second Order Exchangeable Random Fields, page 50

SP Square pyramid, page 104

Bibliography

[1] Khan, S. S., Madden, M. G. (2014). *One-class classification: taxonomy of study and review of techniques.* The Knowledge Engineering Review, 29(3), 345-374.

[2] Pimentel, M. A., Clifton, D. A., Clifton, L., Tarassenko, L. (2014). *A review of novelty detection.* Signal Processing, 99, 215-249.

[3] Khan, S. S., Madden, M. G. (2009, August). *A survey of recent trends in one class classification.* In Irish Conference on Artificial Intelligence and Cognitive Science (pp. 188-197). Springer Berlin Heidelberg.

[4] Tarassenko, L., Hayton, P., Cerneaz, N., Brady, M. (1995). *Novelty detection for the identification of masses in mammograms.*

[5] Kemmler, M., Rodner, E., Wacker, E. S., Denzler, J. (2013). *One-class classification with Gaussian processes.* Pattern Recognition, 46(12), 3507-3518.

[6] Rodner, E., Wacker, E. S., Kemmler, M., Denzler, J. (2011). *One-class classification for anomaly detection in wire ropes with gaussian processes in a few lines of code.* training, 1, 1-5.

[7] Kowalska, K., Peel, L. (2012, July). *Maritime anomaly detection using Gaussian Process active learning.* In Information Fusion (FUSION), 2012 15th International Conference on (pp. 1164-1171). IEEE.

[8] Mery, D., Filbert, D. (2002). *Automated flaw detection in aluminum castings based on the tracking of potential defects in a radioscopic image sequence.* IEEE Transactions on Robotics and Automation, 18(6), 890-901.

[9] Stolfo, S. J., Hershkop, S., Bui, L. H., Ferster, R., Wang, K. (2005, May). *Anomaly detection in computer security and an application to file system accesses.* In Inter-

national Symposium on Methodologies for Intelligent Systems (pp. 14-28). Springer, Berlin, Heidelberg.

[10] Golmohammadi, K., Zaiane, O. R. (2015, October). *Time series contextual anomaly detection for detecting market manipulation in stock market.* In Data Science and Advanced Analytics (DSAA), 2015. 36678 2015. IEEE International Conference on (pp. 1-10). IEEE.

[11] Ayalew, G., Holden, N. M., Grace, P. M., Ward, S. M. (2004). *Detection of glass contamination in horticultural peat with dual-energy X-ray absorptiometry (DXA).* Computers and electronics in agriculture, 42(1), 1-17.

[12] Haff, R. P., Toyofuku, N. (2008). *X-ray detection of defects and contaminants in the food industry.* Sensing and Instrumentation for Food Quality and Safety, 2(4), 262-273.

[13] Patel, D., Davies, E. R., Hannah, I. (1995). *Towards a breakthrough in the detection of contaminants in food products.* Sensor Review, 15(2), 27-28.

[14] Yoon, S. C., Lawrence, K. C., Smith, D. P., Park, B., Windham, W. R. (2007). *Bone fragment detection in chicken breast fillets using transmittance image enhancement.* Transactions of the ASABE, 51(1), 331-339.

[15] Einarsson, G., Jensen, J. N., Paulsen, R. R., Einarsdottir, H., Ersbll, B. K., Dahl, A. B., Christensen, L. B. (2017, June). *Foreign object detection in multispectral x-ray images of food items using sparse discriminant analysis.* In Scandinavian Conference on Image Analysis (pp. 350-361). Springer, Cham.

[16] YangDai, T., Zhang, L. (2016). *Liquid contrabands classification based on energy dispersive X-ray diffraction and hybrid discriminant analysis.* Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 808, 128-134.

[17] Khosa, I., Pasero, E. (2014, January). *Pine nuts selection using X-ray images and logistic regression.* In Computer Applications Research (WSCAR), 2014 World Symposium on (pp. 1-5). IEEE.

[18] Neethirajan, S., Jayas, D. S., White, N. D. G. (2007). *Detection of sprouted wheat kernels using soft X-ray image analysis.* Journal of Food Engineering, 81(3), 509-513.

[19] Casasent, D. A., Sipe, M. A., Schatzki, T. F., Keagy, P. M., Lee, L. C. (1998). *Neural net classification of X-ray pistachio nut data.* LWT-Food Science and Technology, 31(2), 122-128.

[20] Matlab, U. S. G. (1760). *The mathworks.* Inc., Natick, MA, 1992.

[21] Michie, D. J. S. D. (1994). *Machine learning, neural and statistical classification.* Ellis Horwood.

[22] Goldstein, M., Rougier, J. (2006). *Bayes linear calibrated prediction for complex systems.* Journal of the American Statistical Association, 101(475), 1132-1143.

[23] Goldstein, M., Rougier, J. (2009). *Reified Bayesian modelling and inference for physical systems.* Journal of Statistical Planning and Inference, 139(3), 1221-1239.

[24] Goldstein, M., Rougier, J. (2004). *Probabilistic formulations for transferring inferences from mathematical models to physical systems.* SIAM journal on scientific computing, 26(2), 467-487.

[25] McHutchon, A., Rasmussen, C. E. (2011). *Gaussian process training with input noise.* In Advances in Neural Information Processing Systems (pp. 1341-1349).

[26] Maniyar, D. M., Cornford, D., Boukouvalas, A. (2007). *Dimensionality reduction in the emulator setting.* Technical report, Neural Computing Research Group, University of Aston.

[27] Cumming, J. A. and Goldstein, M. (2010) *Bayes linear uncertainty analysis for oil reservoirs based on multiscale computer experiments.* The Oxford handbook of applied Bayesian analysis. Oxford: Oxford University Press, pp. 241-270. Oxford handbooks series.

[28] Goldstein, M., Wooff, D. (2007). *Bayes linear statistics, theory and methods* (Vol. 716). John Wiley Sons.

[29] Brewer, J. (1978). *Kronecker products and matrix calculus in system theory.* IEEE Transactions on circuits and systems, 25(9), 772-781.

[30] Rasmussen, C. E., Williams, C. K. (2006). *Gaussian processes for machine learning* (Vol. 1). Cambridge: MIT press.

[31] Kennedy, M. C., O'Hagan, A. (2001). *Bayesian calibration of computer models.* Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63(3), 425-464.

[32] Vernon, I., Goldstein, M., Bower, R. (2014). *Galaxy formation: Bayesian history matching for the observable universe.* Statistical Science, 29(1), 81-90.

[33] Vernon, I., Goldstein, M., Bower, R. G. (2010). *Rejoinder-Galaxy Formation: a Bayesian uncertainty analysis. Bayesian analysis.*, 5(04), 697-708.

[34] Craig, P. S., Goldstein, M., Rougier, J. C., Seheult, A. H. (2001). *Bayesian forecasting for complex systems using computer simulators.* Journal of the American Statistical Association, 96(454), 717-729.

[35] Craig, P. S., Goldstein, M., Seheult, A. H., Smith, J. A. (1997). *Pressure matching for hydrocarbon reservoirs: a case study in the use of Bayes linear strategies for large computer experiments.* Case studies in Bayesian statistics, 3, 36-93.

[36] Andrianakis, I., Vernon, I. R., McCreesh, N., McKinley, T. J., Oakley, J. E., Nsubuga, R. N., ... White, R. G. (2015). *Bayesian history matching of complex infectious disease models using emulation: a tutorial and a case study on HIV in Uganda.* PLoS computational biology, 11(1), e1003968.

[37] Vernon, I., Liu, J., Goldstein, M., Rowe, J., Topping, J., Lindsey, K. (2016). *Bayesian uncertainty analysis for complex systems biology models: emulation, global parameter searches and evaluation of gene functions.* arXiv preprint arXiv:1607.06358.

[38] Guillas, S., Glover, N., Malki-Epshtein, L. (2014). *Bayesian calibration of the constants of the k turbulence model for a CFD model of street canyon flow.* Computer Methods in Applied Mechanics and Engineering, 279, 536-553.

[39] Goldstein, M., Huntley, N. (2016). *Bayes Linear Emulation, History Matching, and Forecasting for Complex Computer Simulators.* Handbook of Uncertainty Quantification, 1-24.

[40] Rougier, J. *Efficient emulators for multivariate deterministic functions.* Computational and Graphical Statistics, 17 (4): 827843 (2008).

[41] Caiado, C. C. S., Goldstein, M. (2015). *Bayesian uncertainty analysis for complex physical systems modelled by computer simulators with applications to tipping points.* Communications in Nonlinear Science and Numerical Simulation, 26(1), 123-136.

[42] House, L., Goldstein, M., Vernon, I. R. (2009). *Exchangeable computer models.*

[43] Oughton, R. H., Craig, P. S. (2016). *Hierarchical emulation: a method for modeling and comparing nested simulators.* SIAM/ASA Journal on Uncertainty Quantification, 4(1), 495-519.

[44] MUCM (World Wide Web electronic publication, Release 6, 2010). URL: http://www.mucm.ac.uk/.

[45] Nist tables.
`www.nist.gov/pml/data/xraycoef/`

[46] Ibex's MAP technology is covered by UK patents: GB2498615 , GB2532634, GB2532897, GB2533233

[47] Pukelsheim, F. (1994). *The three sigma rule.* The American Statistician, 48(2), 88-91.

[48] Kim, H. M., Mallick, B. K., Holmes, C. C. (2005). *Analyzing nonstationary spatial data using piecewise Gaussian processes.* Journal of the American Statistical Association, 100(470), 653-668.

[49] Gramacy, R. B., Lee, H. K. H. (2008). *Bayesian treed Gaussian process models with an application to computer modeling.* Journal of the American Statistical Association, 103(483), 1119-1130.

[50] De Boor, C., De Boor, C., Mathmaticien, E. U., De Boor, C., De Boor, C. (1978). *A practical guide to splines.* New York: Springer-Verlag.

[51] Goldstein, M. (2013). *Observables and models: Exchangeability and the inductive argument.* Bayesian Theory and Its Applications, 3-18.

[52] Solak, E., Murray-Smith, R., Leithead, W. E., Leith, D. J., Rasmussen, C. E. (2003). *Derivative observations in Gaussian process models of dynamic systems. In Advances in neural information processing systems* (pp. 1057-1064).

[53] Peter A. Parker, G. Geoffrey Vining, Sara R. Wilson, John L. Szarka III, Nels G. Johnson. (2010) *The Prediction Properties of Classical and Inverse Regression for the Simple Linear Calibration Problem. Journal of Quality Technology* 42:4, pages 332-347.

[54] Stephenson, G. (2010). *Using derivative information in the statistical analysis of computer models* (Doctoral dissertation, University of Southampton).

[55] Santner, T. J., Williams, B. J., Notz, W. I. (2013). *The design and analysis of computer experiments.* Springer Science Business Media.

[56] Oakley, J. E., O'Hagan, A. (2004). *Probabilistic sensitivity analysis of complex models: a Bayesian approach.* Journal of the Royal Statistical Society: Series B (Statistical Methodology), 66(3), 751-769.

[57] Kersting, K., Plagemann, C., Pfaff, P., Burgard, W. (2007, June). *Most likely heteroscedastic Gaussian process regression.* In Proceedings of the 24th international conference on Machine learning (pp. 393-400). ACM.

[58] Bastos, L. S., OHagan, A. (2009). *Diagnostics for Gaussian process emulators.* Technometrics, 51(4), 425-438.

[59] Andrianakis, Y. and Challenor, P. G. (2009). Parameter estimation and prediction using Gaussian processes. MUCM Technical report 09/05, University of Southampton.

[60] Saltelli, A., Chan, K. & Scott, E., eds. (2000). *Sensitivity Analysis.* New York: Wiley.

[61] Conti, S., Gosling, J. P., Oakley, J. E., O'Hagan, A. (2009). *Gaussian process emulation of dynamic computer codes.* Biometrika, 96(3), 663-676.

[62] Wang, A., Shapiro, E., Yoon, S., Ganguly, A., Proano, C., Colbeth, R., ... Star-Lack, J. (2015, March). *Asymmetric scatter kernels for software-based scatter correction of gridless mammography.* In SPIE Medical Imaging (pp. 94121I-94121I). International Society for Optics and Photonics.

[63] Siewerdsen, J. H., Daly, M. J., Bakhtiar, B., Moseley, D. J., Richard, S., Keller, H., Jaffray, D. A. (2006). *A simple, direct method for xray scatter estimation and correction in digital radiography and conebeam CT.* Medical physics, 33(1), 187-197.

[64] Wang, L., Shi, L., Wang, D. (2012, January). *Scatter correction in cone-beam CT based on deconvolution.* In Biomedical and Health Informatics (BHI), 2012 IEEE-EMBS International Conference on (pp. 75-78). IEEE.

[65] Maltz, J. S., Blanz, W. E., Hristov, D., Bani-Hashemi, A. (2006, January). *Cone beam X-ray scatter removal via image frequency modulation and filtering.* In Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the (pp. 1854-1857). IEEE.

[66] Aootaphao, S., Thongvigitmanee, S. S., Rajruangrabin, J., Thanasupsombat, C., Srivongsa, T., Thajchayapong, P. (2016). *X-Ray Scatter Correction on Soft Tissue Images for Portable Cone Beam CT.* BioMed research international, 2016.

[67] Ducote, J. L., Molloi, S. (2010). *Scatter correction in digital mammography based on image deconvolution.* Physics in medicine and biology, 55(5), 1295.

[68] Jaenisch, G. R., Ewert, U., Jechow, M. (2010). *Scatter radiation in radiography.* ECNDT, Moscow.

[69] Mentrup, D., Jockel, S., Menser, B., Neitzel, U. (2016). *Iterative scatter correction for grid-less bedside chest radiography: performance for a chest phantom.* Radiation protection dosimetry, 169(1-4), 308-312.

[70] Schrner, K. (2012). *Development of Methods for Scatter Artifact Correction in Industrial X-ray Cone-beam Computed Tomography.* Technical University of Munchen, 131.

[71] Sun, M., Star-Lack, J. M. (2010). *Improved scatter correction using adaptive scatter kernel superposition.* Physics in Medicine Biology, 55(22), 6695.

[72] Seibert, J. A., Boone, J. M. (1988). *Xray scatter removal by deconvolution.* Medical physics, 15(4), 567-575.

[73] Onnia, V., Varjonen, V., Lehtimaki, M., Lehtokangas, M., Saarinen, J. (2000, September). *Scatter compensation in digital radiography.* In Signal Processing Conference, 2000 10th European (pp. 1-4). IEEE.

[74] Zhao, W., Brunner, S., Niu, K., Schafer, S., Royalty, K., Chen, G. H. (2014, March). *A patient-specific scatter artifacts correction method.* In Medical Imaging 2014: Physics

of Medical Imaging (Vol. 9033, p. 903310). International Society for Optics and Photonics.

[75] Chen, Y. W., Han, X., Oikawa, S., Fujita, A. (2007, May). *Independent component analysis for removing x-ray scatter in x-ray images.* In Instrumentation and Measurement Technology Conference Proceedings, 2007. IMTC 2007. IEEE (pp. 1-4). IEEE.

[76] Xie, S. P., Yan, R. J. (2016). *Scattering correction based on regularization deconvolution for Cone-Beam CT.* arXiv preprint arXiv:1604.01581.

[77] Maier, J., Sawall, S., Knaup, M., Kachelrie, M. (2018). *Deep scatter estimation (dse): Accurate real-time scatter estimation for x-ray ct using a deep convolutional neural network.* Journal of Nondestructive Evaluation, 37(3), 57.

[78] Marimn, E., Nait-Charifa, H., Khanb, A., Marsdenc, P. A., Diazd, O. (2017, March). *Scatter reduction for grid-less mammography using the convolution-based image post-processing technique.* In SPIE Medical Imaging (pp. 101324D-101324D). International Society for Optics and Photonics.

[79] Bertero, M., Boccacci, P. (1998). *Introduction to inverse problems in imaging.* CRC press.

[80] El Khettabi, F., Hussein, E. M. (2003). *An inverse problem for three-dimensional x-ray scatter/transmission imaging.* Inverse problems, 19(2), 477.

[81] Jin, K. H., McCann, M. T., Froustey, E., Unser, M. (2017). *Deep convolutional neural network for inverse problems in imaging.* IEEE Transactions on Image Processing, 26(9), 4509-4522.

[82] Wang, L., Mohammad-Djafari, A., Gac, N. (2017, June). *Bayesian X-ray computed tomography using a three-level hierarchical prior model.* In AIP Conference Proceedings (Vol. 1853, No. 1, p. 060003). AIP Publishing.

[83] Siltanen, S., Kolehmainen, V., Jrvenp, S., Kaipio, J. P., Koistinen, P., Lassas, M., ... Somersalo, E. (2003). *Statistical inversion for medical x-ray tomography with few radiographs: I. General theory.* Physics in Medicine Biology, 48(10), 1437.

[84] Singh, V., Jain, A., Bednarek, D. R., Rudin, S. (2014, March). *Limitations of anti-scatter grids when used with high resolution image detectors.* In Proceedings of SPIE–

the International Society for Optical Engineering (Vol. 9033, p. 903362). NIH Public Access.

[85] Rana, R., Jain, A., Shankar, A., Bednarek, D. R., Rudin, S. (2016, March). *Scatter estimation and removal of anti-scatter grid-line artifacts from anthropomorphic head phantom images taken with a high resolution image detector.* In SPIE Medical Imaging (pp. 978364-978364). International Society for Optics and Photonics.

[86] Bushberg, J. T. (Ed.). (2002). *The essential physics of medical imaging.* Lippincott Williams Wilkins.

[87] Salas, S. L., Hille, C. E., Etgen, G. J. (1997). *Calculus: One and several variables.* John Wiley Sons.

[88] Pearson, K. (1916). *Mathematical contributions to the theory of evolution.* XIX. Second supplement to a memoir on skew variation. Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character, 216, 429-457.

[89] Rougier, J., Mangion, A. Z., Schoen, N. (2014). *Computation and Visualisation for large-scale Gaussian updates.* arXiv preprint arXiv:1406.5005.

[90] Khan, M. K., Morigi, S., Reichel, L., Sgallari, F. (2013). *Iterative methods of Richardson-Lucy-type for image deblurring.* Numerical Mathematics: Theory, Methods and Applications, 6(1), 262-275.

[91] Khuri, A. I., Mukhopadhyay, S. (2010). *Response surface methodology.* Wiley Interdisciplinary Reviews: Computational Statistics, 2(2), 128-149.

[92] McDonald, D. B., Grantham, W. J., Tabor, W. L., & Murphy, M. J. (2007). *Global and local optimization using radial basis function response surface models.* Applied Mathematical Modelling, 31(10), 2095-2110.

[93] Pepelyshev, A., Oakley, J. (2009) *On the choice of correlation function and cross-validation for Gaussian processes.* MUCM Technical report 09/07

[94] De Finetti, B. (2017). *Theory of probability: A critical introductory treatment* (Vol. 6). John Wiley Sons.

[95] Goldstein, M. (1986) *Exchangeable belief structures.* Journal of the American Statistical Association 81.396: 971-976.

[96] Kariya, T, Hiroshi K. (2004) *Generalized least squares.* John Wiley & Sons.

[97] Koller, D., Friedman, N. (2009). *Probabilistic graphical models: principles and techniques.* MIT press.

[98] Robert, C. (2007). *The Bayesian choice: from decision-theoretic foundations to computational implementation.* Springer Science Business Media.

[99] Bernardo, J. M. (1996). *The concept of exchangeability and its applications.* Far East Journal of Mathematical Sciences, 4, 111-122.

[100] Goldstein, M. (1997). *Prior inferences for posterior judgements.* In Structures and norms in Science (pp. 55-71). Springer Netherlands.

[101] Goldstein, M. (2006). *Subjective Bayesian analysis: principles and practice.* Bayesian Analysis, 1(3), 403-420.

[102] Berger, J. (2006). *The case for objective Bayesian analysis.* Bayesian analysis, 1(3), 385-402.

[103] Williamson, D., Goldstein, M. (2015). *Posterior belief assessment: Extracting meaningful subjective judgements from Bayesian analyses with complex statistical models.* Bayesian Analysis, 10(4), 877-908.

[104] Bernardo, J. M., Bayarri, M. J., Berger, J. O., Dawid, A. P., Heckerman, D., Smith, A. F. M., West, M. (2011). *External Bayesian analysis for computer simulators.* Bayesian Statistics 9, 9, 201.

[105] Rourgier, J. (2014) *https://www2.warwick.ac.uk/fac/sci/statistics/apts/* Course in Statistical Inference.

[106] Mahalanobis, P. C. (1936). *On the generalised distance in statistics.* Proceedings of the National Institute of Sciences of India, 1936, 49-55.

[107] Hosmer Jr, D. W., Lemeshow, S., Sturdivant, R. X. (2013). *Applied logistic regression (Vol. 398).* John Wiley Sons.

[108] Currin, C. (1988). *A Bayesian approach to the design and analysis of computer experiments* (No. ORNL-6498). ORNL Oak Ridge National Laboratory (US).

[109] Bor, D., Birgul, O., Onal, U., Olgar, T. (2016). *Investigation of grid performance using simple image quality tests.* Journal of Medical Physics / Association of Medical Physicists of India, 41(1), 2128. http://doi.org/10.4103/0971-6203.177280

[110] Belykh, I. (2014). Grid Artifacts Suppression in Computed Radiographic Images. World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering, 8(8), 1402-1405.

[111] Correa, S. C., Souza, E. M., Silva, A. X., Lopes, R. T. *Modeling and analysis of anti scatter grids for analogical and digital systems using MCNPX.*

[112] Bernardo, J. M. (1996). The concept of exchangeability and its applications. Far East Journal of Mathematical Sciences, 4, 111-122.

[113] Agostinelli, S., Allison, J., Amako, K. A., Apostolakis, J., Araujo, H., Arce, P., ... Behner, F. (2003). *GEANT4a simulation toolkit.* Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 506(3), 250-303.

[114] Vernon, I., Gosling, J. P. (2017). *A Bayesian computer model analysis of robust Bayesian analyses.* arXiv preprint arXiv:1703.01234.

[115] Rougier, J. (2007). *Probabilistic inference for future climate using an ensemble of climate model evaluations.* Climatic Change, 81(3-4), 247-264.

[116] Solak, E., Murray-Smith, R., Leithead, W. E., Leith, D. J., Rasmussen, C. E. (2003). *Derivative observations in Gaussian process models of dynamic systems.* In Advances in neural information processing systems (pp. 1057-1064).

[117] Jones, M., Goldstein, M., Jonathan, P., Randell, D. (2016). *Bayes linear analysis for Bayesian optimal experimental design.* Journal of Statistical Planning and Inference, 171, 115-129.

[118] Holden, P. B., Edwards, N. R., Hensman, J., Wilkinson, R. D. (2015). *ABC for climate: dealing with expensive simulators.* arXiv preprint arXiv:1511.03475.

[119] Higham, N. J. (1990). *Analysis of the Cholesky decomposition of a semi-definite matrix* (pp. 161-185). Oxford University Press.

[120] Mola, F., Conversano, C., Vichi, M. (Eds.). (2018). *Classification,(Big) Data Analysis and Statistical Learning.* Springer International Publishing.

[121] Flower, M. A. (Ed.). (2012). *Webb's physics of medical imaging.* CRC Press.

[122] Guthoff, M., Brovchenko, O., de Boer, W., Dierlamm, A., Mller, T., Ritter, A., ... Simonis, H. J. (2012). *Geant4 simulation of a filtered X-ray source for radiation damage studies* Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 675, 118-122.

[123] Boone, J. M., Cooper, V. N. (2000). *Scatter/primary in mammography: Monte Carlo validation.* Medical physics, 27(8), 1818-1831.

[124] Leon, S. M., Brateman, L. F., Wagner, L. K. (2014). *Characterization of scatter in digital mammography from use of Monte Carlo simulations and comparison to physical measurements.* Medical physics, 41(11).

[125] Hamermesh, M. (1962). *Group theory.* Reading, Mass, 368.

Appendix A

Appendix One

A.1 Notation for Multivariate Random Fields

In this thesis, it is often be necessary to i) reformulate matrices as vectors, ii) reference subsets of matrices and iii) reference sets of locations on multivariate random fields. The notation was chosen to align with the computing language in which the majority of the methods were coded.

1. The vectorisation of $(n \times m)$ matrix x is

$$x_{(:)} = (x_{(:,1)}^T, \cdots, x_{(:,m)}^T)^T \tag{1.1.1}$$

   where $x_{(i,:)}$ is column i of x.

2. If a set of locations is notated $T' = (t'^{(1)}, \cdots, t'^{(\tilde{n})})^T$, then that set of locations on the multivariate random field is notated

$$y(T') = (y(t'^{(1)}), y(t'^{(2)}), \cdots y(t'^{(\tilde{n})}))^T \tag{1.1.2}$$

   where $y(T')$ is a $(\tilde{n} \times q)$ matrix. Every location on a field is notated $y = y(T)$.

3. A set of locations on a constituent field is notated

$$y_{(k)}(T') = (y_{(k)}(t'^{(1)}), y_{(k)}(t'^{(2)}), \cdots y_{(k)}(t'^{(\tilde{n})}))^T \tag{1.1.3}$$

   where $y_{(k)}(T')$ is a $(\tilde{n} \times 1)$ vector.

4. A vectorised set of locations of a multivariate random field is notated

$$y_{(:)}(T') = \Big(y_{(1)}(t'^{(1)}), \cdots, y_{(1)}(t'^{(\tilde{n})}), y_{(2)}(t'^{(1)}), \cdots, y_{(2)}(t'^{(\tilde{n})}), \cdots, y_{(q)}(t'^{(1)}), \cdots, y_{(q)}(t'^{(\tilde{n})})\Big)^T \tag{1.1.4}$$

where $y_{(:)}(T')$ is a $(\tilde{n}q \times 1)$ vector.

5. If x is a $(n \times m \times p)$ matrix, then the $x_{(:,k)}$ vectorises the first two dimensions, i.e.

$$x_{(:,k)} = a_{(:)} \text{ where } a = x_{(:,:,k)}. \tag{1.1.5}$$

6. Unless otherwise stated, when beliefs are adjusted using a matrix, it is assumed to be vectorised i.e.

$$\mathbb{E}_z[x] = \mathbb{E}_{z_{(:)}}[x_{(:)}],\ \mathbb{V}ar_z[x] = \mathbb{V}ar_{z_{(:)}}[x_{(:)}] \text{ and } \mathbb{C}ov_z[x,y] = \mathbb{C}ov_{z_{(:)}}[x_{(:)}, y_{(:)}]. \tag{1.1.6}$$

7. Sets of quantities are notated by

$$\begin{aligned}&\left\{x_{(i)}^{(j)}(t^{(k)})\right\}_{i=1:n,j=1:n',k=1:n''} = \\ &\left(x_{(1)}^{(1)}(t^{(1)}), \cdots, x_{(n)}^{(1)}(t^{(1)}), x_{(1)}^{(2)}(t^{(1)}), \cdots, x_{(n)}^{(2)}(t^{(1)}), \cdots, x_{(n)}^{(n')}(t^{(1)}), \cdots, x_{(n)}^{(n')}(t^{(n'')}\right).\end{aligned} \tag{1.1.7}$$

A.2 Splines

A.2.1 Truncated Polynomials

An approach to specifying a piece-wise polynomial is a truncated polynomial basis.

Definition A.2.1. ***Truncated Polynomial Basis.*** *Given a set of knots* $\{k_{(1)}, \cdots, k_{(l)}\}$, *the truncated polynomial spline basis function of degree* $s > 0$ *is*

$$h(t) = (1, t^1, \cdots, t^{s-1}, (t - k_{(1)})_+^s, \cdots, (t - k_{(l-1)})_+^s) \text{ where} \tag{1.2.1}$$

$$(t)_+^s) = t^r \ for t > 0 \text{ and } 0 \text{ otherwise.} \tag{1.2.2}$$

Truncated polynomials are logical and useful but suffer from numerical instability as the matrix $H^T H$ is often badly conditioned.

A.2.2 B-splines

A numerically more stable alternative is a B-spline.

Definition A.2.2. ***B-Spline Basis.*** *Given a set of notes* $\{k_{(1)}, \cdots, k_{(l)}\}$*, the B-spline basis of degree* $s > 0$ *is given by*

$$h(t) = (\mathcal{B}(t, s, 1), , \cdots, \mathcal{B}(t, s, l + s - 1)) \text{ where} \tag{1.2.3}$$

$$\mathcal{B}(t, s, j) = \frac{t - k_{(j-s)}}{k_{(j)} - k_{(j-r)}}\mathcal{B}(t, s-1, j-1) + \frac{k_{(j+1)} - t}{k_{(j+1)} - k_{(j+1-s)}}\mathcal{B}(t, s-1, j) \text{ and} \tag{1.2.4}$$

$$\begin{aligned} &\mathcal{B}(t, 0, j) = 1 \text{ for } k_{(j)} \leq t < k_{(j+1)} \text{ and} \\ &\mathcal{B}(t, 0, j) = 0 \text{ otherwise.} \end{aligned} \tag{1.2.5}$$

B-splines are a good choice as they are usually better conditioned than, for example, truncated polynomials. The residual component may be specified as continuous or independent by adjusting the covariance terms in $\mathbb{V}ar[\mathcal{W}(T)]$.

A.3 Automated Variance Specification

Due to the difficulties in elicitation, it is of interest to consider an automated process for fourth order specifications requiring only the second order specification of more reasonable random vectors. The specifications of the fourth order quantities are then produced using sampling experiments. The method assumes that the $R^{(i)}(y)$ may be written in the form

$$R^{(i)}(y_{(:)}) = LZ^{(i)}_{(:)} \text{ where } LL^T = M(V). \tag{1.3.1}$$

Here $\mathcal{M}(V) \perp\!\!\!\perp Z^{(i)}$, $\mathbb{E}[Z^{(i)}] = 0$ and $\mathbb{V}ar[Z^{(i)}] = I_q$. For this automated specification, a number of specifications are required: i) a second order specification for the diagonal elements of $\mathcal{M}(V)$ and the non-diagonal elements of the correlation matrix and ii) a distribution for Z, $\mathcal{M}(V)$ and the elements of the correlation matrix. The sampling algorithm is as follows:

1. sample prior correlations for the population variance matrix, $\text{Corr}(M(V_{(k,j)}))$ for $k < j$,
2. sample prior population variances $\text{diag}(M(V))$,
3. calculate prior population variance matrix and calculate L using a cholesky decomposition,
4. sample a data set size N and
5. calculate the sample variance estimator.

This is repeated until the uncertainty in $\mathbb{V}ar[\hat{\Sigma}]$ and $\mathbb{V}ar[M(V)]$ has been sufficiently resolved. Figure A.1 shows outputs of the automated fourth order specification method, where all distributions were specified as uniform, for a random vector with two dimensions. From these outputs, a fourth order specification can be made using the usual unbiased estimators for the population mean, variance and covariance.

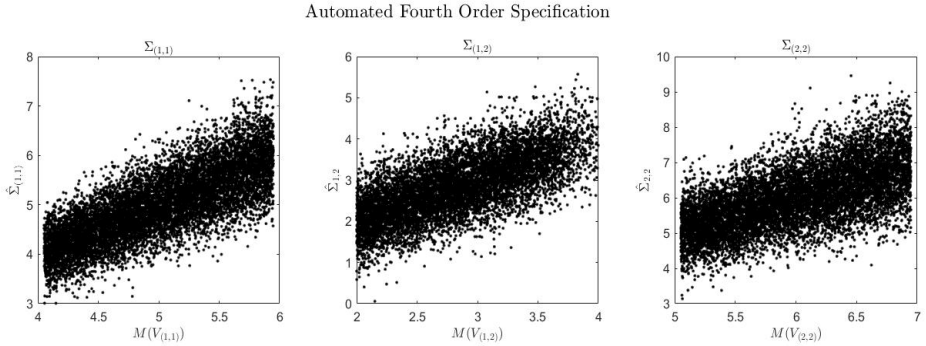

Figure A.1: Left: plot of population variance against estimator of population variance. Centre: plot of population covariance against estimator for population covariance. Right: plot of population variance against estimator of population variance.

Specifying expectations and variances, of population variances and correlations, is a more reasonable expectation for expert specification. However, full probabilistic specifications are required in order to perform sampling experiments. If an expert has to make these specifications, a full Bayes analysis could be performed. However, full Bayes requires a large computational outlay every time inference is performed; this can be impractical particularly for online algorithms. On the other hand, the automated variance specification can be performed offline retaining efficiency for online inference.

A.3.1 Robust Variance Learning

An expert has given a second order specification and sampling experiments have been used to make a fourth order specification. In order to perform the sampling experiments, distributions are fixed and these distributions have an effect on the specifications. The expert has not asserted that these distributions are their probabilities so it is of interest to consider what the impact is if the fixed distributions are different to those the expert would assert given they had more time. This is a type of sensitivity analysis which can inform i) whether additional effort is likely to have a significant impact on the inference, ii) the best place to allocate resources and iii) whether certain specifications are robust against

expensive errors. Here methods for assessing two features will be presented: the first is assessing the sensitivity of the fourth order specification to the choice of distribution and the second is investigating how fixing the distribution affects the second order properties of the adjusted version.

Sensitivity of Fourth Order Specification

The method considers a functional relationship between i) the distributions specified and ii) the outputs of the automated fourth order specification. The relationship is given by

$$\mathcal{G}(x_{ps}) = y_{FOS} \tag{1.3.2}$$

where $\mathcal{G}$ describes the automated specification algorithm, x_{ps} are parameters describing the distributional assumptions and y_{FOS} is the fourth order specification. To be able to use the form given in equation 1.3.2, a discrete parameterisation of the distributional assumptions is required. One way of creating a discrete parameterisation is when Z is assumed to be a linear combination of independent random quantities

$$Z = \frac{1}{a}\sum_{i=1}^{b} \alpha_{(i)} Z_{(i)} \tag{1.3.3}$$

where $\sum_{i=1}^{b} \alpha_{(i)} = 1$, $\mathbb{E}[Z_{(i)}] = 0$, $\mathbb{V}ar[Z_{(i)}] = I$ and a is a normalising constant to ensure that the $\mathbb{V}ar[Z] = I$. This construction is advantageous because it shares an analogous formalism to the analysis of complex computer models. This form means the tools for sensitivity and uncertainty analysis of simulators can be applied to robustness analysis. If the number of parameters (required to describe the distributional specification) is large and the sampling algorithm computationally expensive, a statistical model of $\mathcal{G}$ may be built and the emulation approach (see chapter 4) to the analysis of complex computer models applied [114].

The approach is exemplified by an investigation into the sensitivity of fourth order specifications to distributional assumptions. The random quantities in the linear combination are defined by

$$\begin{aligned} Z_{(1)} &\sim \mathcal{N}(0,1) \\ Z_{(2)} &\sim \mathcal{U}(0-\sqrt{3}, 0+\sqrt{3}) \\ Z_{(3)} &\sim \mathcal{SN}(0,1,-1) \end{aligned} \tag{1.3.4}$$

where $\mathcal{SN}$ is a skew normal distribution, sampled from the Pearson family of distributions with skewness minus one and kurtosis three [88]. The function to be analysed is

$$\mathcal{G}(\alpha) = (\mathbb{V}ar[M(V_{(:)})], \mathbb{V}ar[\hat{\Sigma}_{(:)}]). \tag{1.3.5}$$

Here $(\mathbb{V}ar[M(V)], \mathbb{V}ar[\hat{\Sigma}])$ is the output of an automated fourth order specification. Figure A.2 shows $\mathcal{G}_{(1)}(\alpha) = \mathbb{V}ar[\hat{\Sigma}_{(1,1)}]$ for a range of alpha. The plots show that decreasing the kurtosis decreases the estimator uncertainty and skewness has minimal impact on the specification. Therefore, if there are resources available to decrease the impact of fixing the distribution, they are best spent on resolving uncertainty in the kurtosis of y.

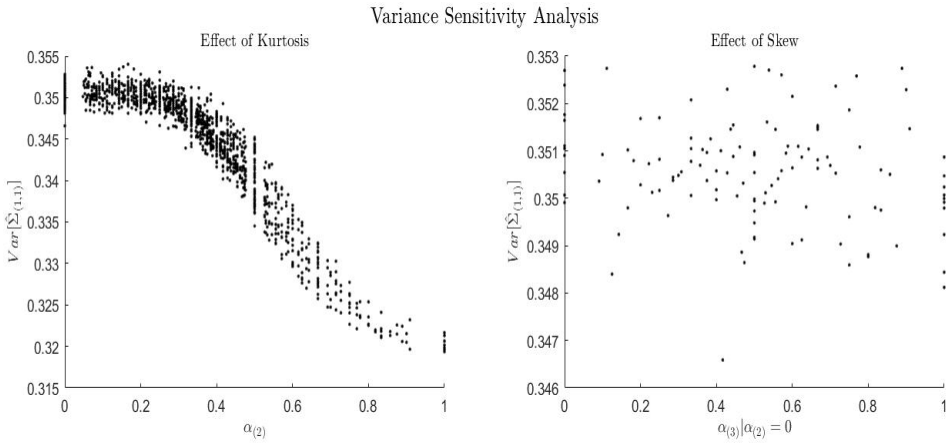

Figure A.2: Left: variance of informative statistic against proportion of uniform distribution in mixture. Right: variance of informative statistic against proportion of skew normal in mixture.

Sensitivity of Adjusted Versions

Another pertinent question is, if there is discrepancy between the distributional specification and the actual distribution, what is the impact on the first and second central moments of the adjusted versions? This time the functional relationship is

$$\mathcal{G}'(x_{ps}, x_{ss}) = y_{AV}. \tag{1.3.6}$$

Here $\mathcal{G}'$ describes a sampling algorithm which produces a forth order specification dependent on parameters x_{ps} and calculates the second order statistics of the adjusted versions, from observations sampled from distributions defined by x_{ss}. The subscripts i) on x_{ps} notate prior specification, ii) on x_{ss} sampling specification and iii) on y_{AV} adjusted versions.

The sensitivity of the adjustment to fixing the sampling distributions is illustrated using

$$\mathcal{G}'(\alpha_p, \alpha_t) = (\bar{\mathbb{E}}[\mathbb{A}_{Df}[M(V)]], \bar{\mathbb{V}ar}[\mathbb{A}_{Df}[M(V)]]). \tag{1.3.7}$$

Here α_p and α_t are parameters as in equation 1.3.3. $\mathcal{G}'$ is a sampling algorithm that i) produces a fourth order specification using a distribution defined by α_p, ii) samples 1000 sets of 20 observations using a distribution defined by α_t, iii) calculates 1000 adjusted versions and iv) calculates the sample mean and variance of those adjusted versions. In this case, the distributions were

$$\begin{aligned} Z_{(1)} &\sim \mathcal{N}(0,1) \text{ and} \\ Z_{(1)} &\sim \mathcal{U}(0-\sqrt{3}, 0+\sqrt{3}). \end{aligned} \tag{1.3.8}$$

Then, for example, the distribution used to make the fourth order specification is sampled using

$$Z_{(p)} = \frac{1}{a}(\alpha_p Z_{(2)} + (1-\alpha_p)Z_{(2)}). \tag{1.3.9}$$

This example assesses the difference $\bar{\mathbb{V}ar}[\mathbb{A}_{D^f}[M(V)]] - \mathbb{V}ar[\mathbb{A}_{D^f}[M(V)]]$, which is the discrepancy between an estimator for the adjusted variance and the prior specification for the adjusted variance. In other words, beliefs are adjusted assuming the distribution is given by α_p when in fact it is given by α_t.

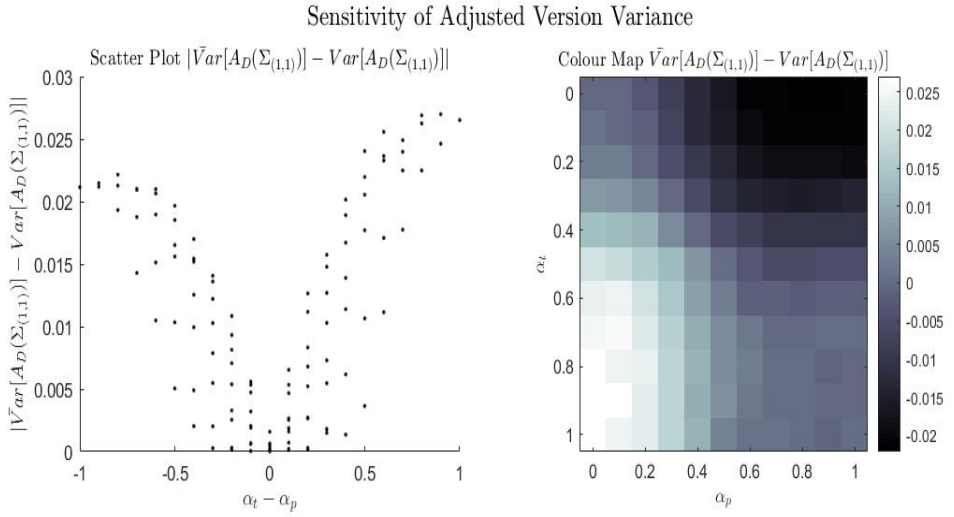

Figure A.3: Left: scatter plot of difference between prior specification and sampling distribution against absolute difference between specified adjusted version variance and actual adjusted version variance. Right: colour map of difference between specified adjusted version variance and actual adjusted version variance.

Figure A.3 shows analysis of the uncertainty estimation for the illustrative example. When $\alpha_t = \alpha_p$, the expected value of $\bar{\mathbb{V}ar}[\mathbb{A}_{D^f}[M(V)]] - \mathbb{V}ar[\mathbb{A}_{D^f}[M(V)]]$ is zero. As $|\alpha_t - \alpha_p|$ increases so does $|\bar{\mathbb{V}ar}[\mathbb{A}_{D^f}[M(V)]] - \mathbb{V}ar[\mathbb{A}_{D^f}[M(V)]]|$. This analysis enables

us to choose specifications which are robust to certain errors. For example, if an expert does not want to underestimate adjusted variance, they would then choose an α_p for which the difference $\bar{\mathbb{V}ar}[\mathbb{A}_{Df}[M(V)]] - \mathbb{V}ar[\mathbb{A}_{Df}[M(V)]]$ is most often negative; in this example, when $\alpha_p = 1$. Variance learning is challenging as it is difficult to elicit beliefs about the quantities involved. This section has presented an algorithm for specifying these quantities requiring a more reasonable second order specification. This algorithm also requires distributional specifications which an expert may not be able to provide. In order to ensure that these specifications are not having a large impact on the inference, robustness analysis is performed, constructed in such a way that the methodologies for analysing complex computer models can be applied.

A.4 Variance Learning SOERFs

In this section, a set of simplifying assumptions and approximate methods for variance learning, for second order exchangeable random fields, are presented.

A.4.1 Automated Sampling Method

This sampling approach is general and will utilise the prior specifications that the expert has made.

1. Sample the diagonal elements of $M(V_{\mathcal{W}_{(k)}(T)})$ and $M(V_{\beta_{(:,:,k)}})$.
2. Sample the correlation fields $M(C_{(k,k')})$.
3. Sample the correlation field $M(C_{(k,k')}(T))$ for all $p < p'$.
4. Using the outputs of steps 1), 2) and 3), create a sample of $M(V_{\mathcal{W}_{(k)}(t)})$ and $M(V_{\beta_{(k)}})$. The prior correlation matrix $\mathrm{Corr}(M(V_{\mathcal{W}_{(k,k')}(T)}))$ is fixed for all $k < k'$.
5. Calculate the Cholesky decomposition of the samples $M(V_{\mathcal{W}_{(k)}(t)})$ and $M(V_{\beta_{(k)}})$
6. Sample the population means $M(\mathcal{W}_{(k)})$ and $M(\beta_{(k)})$.
7. Using the outputs of 5) and 6), sample N observations.
8. Calculate $\hat{\Sigma}_{\beta_{(k)}}$ and $\hat{\Sigma}_{\mathcal{W}_{(k)}}(T)$.
9. Repeat until the uncertainty in the expectation and variance of the estimators has been well resolved.

This method requires the second order specification of the i) diagonal elements of $M(V_{\mathcal{W}_{(k)}(T)})$ and $M(V_{\beta_{(k)}})$ ii) the correlation field $M(C_{(k,k')}(T))$ and iii) non-diagonal elements of the correlation matrices of $M(V_{\beta_{(k)}})$. A choice of distribution has to be made in order to sample the data; for the fields, Gaussian process assumptions are a good starting place. The problem is that, as distributions have been specified, which are not asserted by the expert, additional work has to be done to ensure the inference is robust to those specifications. This process is likely to be expensive, especially for large fields.

A.4.2 Direct Adjustment

The direct approach to adjustment follows from [28] chapter 8. The idea is to calculate the adjusted beliefs as a weighted averaged of prior beliefs and informative statistics. The adjusted expectation of $M(V_{\beta(k)})$ by $\hat{\Sigma}_{\beta_{(k)}}$ is given by

$$\mathbb{E}_{\hat{\Sigma}_{\beta_{(k)}}}[M(V_{\beta(k)})] = (1-\alpha)\mathbb{E}[M(V_{\beta(k)})] + \alpha(\hat{\Sigma}_{\beta_{(k)}} - \mathbb{E}[T_{\beta_{(k)}}]). \tag{1.4.1}$$

Here $\alpha = \frac{N}{N+m}$ where m is the notional sample size attached to the prior specification. For the variance component another step is added to the direct adjustment. The direct adjustment is calculated at every location on the field. Then the adjusted expectation is calculated using

$$\mathbb{E}_{\hat{\Sigma}_{\mathcal{W}_{(k)}}(t)}[M(V_{\mathcal{W}_{(k)}(t)})] = \sum_{t' \in T} W(t') \left((1-\alpha)\mathbb{E}[M(V_{\mathcal{W}_{(k)}(t')})] + \alpha(\hat{\Sigma}_{\mathcal{W}_{(k)}}(t') - \mathbb{E}[T_{\hat{\Sigma}_{\mathcal{W}_{(k)}}(t')}]) \right) \tag{1.4.2}$$

where W is a weight matrix which described the relative notional sample sizes in t. A good method of designing W is to use

$$W = \frac{1}{a} C(t', T; l)(C(T, T; l) + \sigma_\epsilon I_n)^{-1} \tag{1.4.3}$$

where σ_ϵ is specified directly, C is a correlation function, l is a correlation length and a normalises the W so that $\sum_i W_{(t^{(i)})} = 1$. σ_ϵ is the ratio of uncertainty in the population variance and the estimators squared standardised error. This is a direct adjustment which utilises prior expectation and then a smoother which utilises spatial correlation.

A.4.3 Adjusted Expectation Variance Learning

Another option is to adjust beliefs for the population mean and variance of the adjusted expectations. The set of adjusted expectations for the regression parameters are second

order exchangeable and the set of adjusted expectations of the residual process are second order exchangeable random fields. This implies that

$$\mathbb{E}_{z^{(i)}}[\beta^{(i)}_{(k)}] = M(\mathbb{E}_z[\beta_{(k)}]) \oplus R^{(i)}(\mathbb{E}_z[\beta_{(k)}]) \text{ and} \tag{1.4.4}$$

$$\mathbb{E}_{z^{(i)}}[\mathcal{W}^{(i)}_{(k)}(t)] = M(\mathbb{E}_z[\mathcal{W}_{(k)}(t)]) \oplus R^{(i)}(\mathbb{E}_z[\mathcal{W}_{(k)}(t)]) \tag{1.4.5}$$

and, if second order exchangeability assumptions are made over the appropriate squared quantities,

$$R^{(i)}(\mathbb{E}_z[\beta_{(k)}])R^{(i)}(\mathbb{E}_z[\beta_{(k)}])^T = M(V_{R(\mathbb{E}_z[\beta_{(k)}])}) \oplus R^{(i)}(V_{R(\mathbb{E}_z[\beta_{(k)}])}) \text{ and} \tag{1.4.6}$$

$$R^{(i)}(\mathbb{E}_z[\mathcal{W}_{(k)}(t)])R^{(i)}(\mathbb{E}_z[\mathcal{W}_{(k)}(t)])^T = M(V_{R(\mathbb{E}_z[\mathcal{W}_{(k)}(t)])}) \oplus R^{(i)}(V_{R(\mathbb{E}_z[\mathcal{W}_{(k)}(t)])}). \tag{1.4.7}$$

Then variance learning methods can be applied to these quantities without the issue that both the quantity and population mean are unknown. These quantities are not the population mean and variance of the second order exchangeable set of multivariate random fields. Therefore, this approach should only be used if $M(V_{R(\mathbb{E}_z[\beta_{(k)}])})$ and $M(V_{R(\mathbb{E}_z[\mathcal{W}_{(k)}(t)])})$ are sufficient to perform the inferential tasks required.

A.4.4 Anomaly Detection with Relative Population Size and the Anomalous Population

Other beliefs, that have an impact on the inference, may be available; for example, the relative size of the populations. In the full Bayesian approach, this is easily incorporated. However, with Bayes linear, it is more challenging. This is because linear fitting does not work well for the belief adjustment of indicator functions. Indeed, an adjusted expectation of an indicator function is not a probability. An approach to incorporating all the necessary information is to use

$$\frac{Dis(z; A=1)^{-1}\mathbb{P}(A=1)}{Dis(z; A=0)^{-1}\mathbb{P}(A=0) + Dis(z; A=1)^{-1}\mathbb{P}(A=1)} \tag{1.4.8}$$

where

$$Dis(z; A=1) = (z - \mathbb{E}[z; A=1])^T \mathbb{V}ar[z; A=1]^{-1}(z - \mathbb{E}[z; A=1]) \text{ and} \tag{1.4.9}$$

$$Dis(z; A=0) = (z - \mathbb{E}_{D^f}[z; A=0])^T \mathbb{V}ar_{D^f}[z; A=0]^{-1}(z - \mathbb{E}_{D^f}[z; A=0]). \tag{1.4.10}$$

This measure belongs to $[0, 1]$, it includes beliefs about the relative population size and judgements about the distribution of the anomalous population. Large values indicate the sample is anomalous and small values non-anomalous. However, it is not an expert's

probability. The advantage of $\mathbb{P}(A = 1; z)$ is that it is clear how to proceed. For example, if an expert aims to wrongly identify non-anomalous as anomalous once in every one hundred observations, the decision threshold is set as $\mathbb{P}(A = 1|z) \geq 0.99$. With $Dis(z)$ this task is more challenging and it is not clear exactly what $Dis(z)$ means. However, if the full Bayesian approach uses approximations and simplifications, the same criticism can be levelled at $\mathbb{P}(A = 1|z)$. The probabilities specified are encoding no individuals posterior probability and $\mathbb{P}(A = 1|z)$ is also a measure summarising the specifications that were made and some more arbitrary judgements. The approach to inference presented in this section does not sit on the comforting coherent foundations of the full Bayesian approach. In practice however, it is often similar and less expensive.

A.5 Processing the MAP

This section will give further details on how continuous energy parameters are extracted from X-ray images with a MAP. As the method must be efficient and minimise loss of spatial resolution, instead of using ordinary least squares, the weighted least squares approach is preferred. The new estimates are given by

$$\beta^{EC}(i,j) = (H^T W(i,j) H)^{-1} H^T W(i,j) \log(\mathcal{T}) \tag{1.5.1}$$

where H is a design matrix for the continuous energy model and $W(i,j)$ is a weight matrix. Given I and I_0, the continuous energy regression parameter at pixel (i,j) is given by

$$\begin{aligned}
\beta_1^{EC}(i,j) &= \frac{A}{B} \\
A &= \sum_{l=1}^{\kappa}\sum_{k=1}^{\theta} W_{(l,k)}(i,j)\log(\mathcal{T}(l,k))I_0(l,k) - \\
&\quad \sum_{l=1}^{\kappa}\sum_{k=1}^{\theta} W_{(l,k)}(i,j)\log(\mathcal{T}(l,k))\sum_{l=1}^{\kappa}\sum_{k=1}^{\theta} W_{(l,k)}(i,j)I_0(l,k) \\
B &= \sum_{l=1}^{\kappa}\sum_{k=1}^{\theta} W_{(l,k)}(i,j)(I_0(l,k))^2 - \left(\sum_{l=1}^{\kappa}\sum_{k=1}^{\theta} W_{(l,k)}(i,j)I_0(l,k)\right)^2 \\
\beta_0^{EC}(i,j) &= \sum_{l=1}^{\kappa}\sum_{k=1}^{\theta} W_{(l,k)}(i,j)\log(\mathcal{T}(l,k)) - \sum_{l=1}^{\kappa}\sum_{k=1}^{\theta} W_{(l,k)}(i,j)I_0(l,k)\frac{A}{B}.
\end{aligned} \tag{1.5.2}$$

If W can be assumed to be approximately stationary, each individual term may be calculated using the fast Fourier transform and the convolution theorem.

Functional Weights

The simplest way to specify a weight is a functional kernel, for example a Gaussian or moving average filter. The advantage is that the response of the system is predictable and calculations are simple. The disadvantage is that i) no information about edges is incorporated implying sub-optimal spatial resolution and ii) a correlated noise structure is exacerbated. An alternative, which has better spatial resolution and does not significantly exacerbate the correlated noise structure, is a Fourier notching approximation.

Fourier Notching Approximations

Instead of picking the weights with a functional kernel, an approximate Fourier based filtering technique can be utilised. This technique – often referred to as notch filtering – attenuates particular frequencies in the Fourier domain. Given the oscillating structure of a MAP, the Fourier transform of an image will exhibit distinctive peaks around known frequencies. These peaks may be detected and a mask created which is an image with ones everywhere and zeros at those peaks. The Fourier transform of the image is multiplied by the mask and then transformed back to the space domain. For a less aggressive smoother, a notching function may be utilised which is zero at the maxima of the peaks and smoothly increases to one. In this work, a Dolph-Chebyshev window is used. The Dolph-Chebyshev has theoretical advantages but testing showed that the choice of windowing function had a small impact on the output.

The notching technique produces an estimate for a local average with higher spatial resolution. However, as the space domain response of the mask need not be positive, the process does not produce a valid weight matrix. The notching technique can therefore only be used to calculate a higher resolution estimate of the components

$$\sum_{l=1}^{\kappa}\sum_{k=1}^{\theta} W_{(l,k)}(i,j)\log(\mathcal{T}(l,k)) \text{ and} \tag{1.5.3}$$

$$\sum_{l=1}^{\kappa}\sum_{k=1}^{\theta} W_{(l,k)}(i,j)I_0(l,k). \tag{1.5.4}$$

The notching technique is a higher resolution estimate of the local sample mean, which is approximately unbiased in smooth regions, but not over edges. Figure A.4 shows a mask and the equivalent weight matrix. The weights do not meet the assumptions of weighted least squares as they are not all positive.

Figure A.5 shows a comparison of the local means using the two methods for creating a

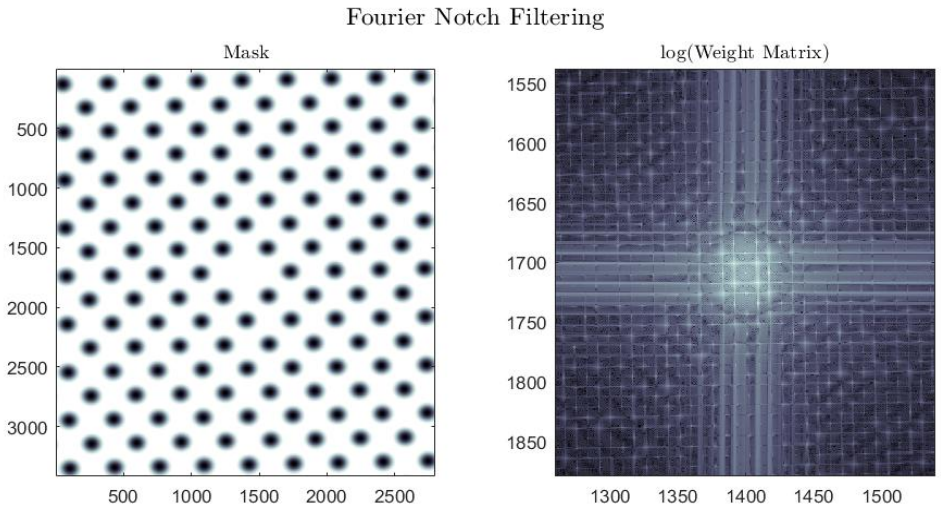

Figure A.4: Left: mask used to attenuate the peaks induced by MAP in frequency space. Right: equivalent weight matrix computed using the inverse fast Fourier transform of the mask.

weight matrix. The notching method has higher spatial resolution due to the lower degree of smoothing. The difference plots show that for flat regions the expected difference between the images is approximately zero. This illustrates why the Fourier weights can be used for the first order moments in the generalised least squares calculations. Using the notching approach for the first order moments means that the β^{EC} has better resolution than a Gaussian weight.

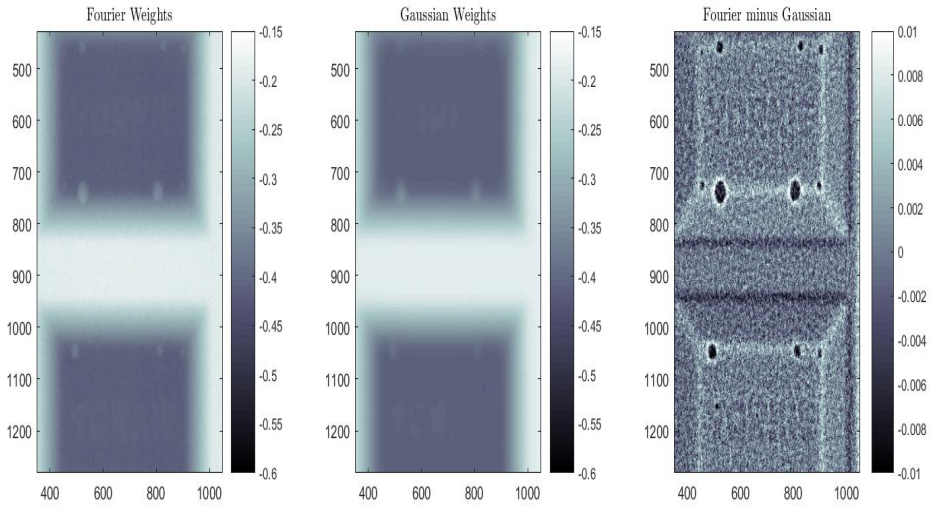

Figure A.5: Left: local mean of an image of a chocolate bar calculated with Fourier notching. Centre: local mean calculated with Gaussian weights. Right: difference between the two methods of calculating a local mean.

A.6 Bayes Linear Emulator of a Stochastic Simulator

In this section, a more complete approach to modelling stochastic simulators will be presented. This method was tested in early versions of the algorithm presented in section 4.3.4. In this case, it is the simulator evaluated at a set of points, plus repetitions

$$\{X, D^{\mathcal{F}}\} = \begin{bmatrix} x^{(1)} & , & \mathcal{F}^{(1)}(x^{(1)}), \cdots, \mathcal{F}^{(\mathbb{N}_1)}(x^{(1)}) \\ \vdots & & \vdots \\ x^{(\mathbb{N})} & , & \mathcal{F}^{(1)}(x^{(\mathbb{N})}), \cdots, \mathcal{F}^{(\mathbb{N}_{\mathbb{N}})}(x^{(\mathbb{N})}) \end{bmatrix}.$$

The emulator is then defined in two stages: the first emulator organises beliefs about the simulator's population mean and the second organises beliefs about its variance i.e.

$$\begin{aligned} M(\mathcal{F}_{(t)}(x)) &= \mathcal{M}_{M(t)}(x) \oplus \mathcal{W}_{M(t)}(x) \text{ and} \\ M(V_{\mathcal{F}(t)}(x)) &= \mathcal{M}_{V(t)}(x) \oplus \mathcal{W}_{V(t)}(x). \end{aligned} \tag{1.6.2}$$

Beliefs about the population mean and variance of the simulator are adjusted using the unbiased estimators for the population mean and variance, given by

$$\{ \text{ X, D}^{\bar{\mathcal{F}}}\} = \left\{x^{(i)}, \bar{\mathcal{F}}(x^{(i)}), \hat{\Sigma}_{\mathcal{F}}(x^{(i)})\right\}_{i=1:\mathbb{N}} \tag{1.6.3}$$

where

$$\bar{\mathcal{F}}(x^{(i)}) = \frac{1}{\mathbb{N}_i} \sum_{j=1}^{\mathbb{N}_i} \mathcal{F}^{(j)}(x^{(i)}) \text{ and} \tag{1.6.4}$$

$$\begin{aligned} \hat{\Sigma}_{\mathcal{F}}(x^{(i)}) &= \frac{1}{\mathbb{N}_i - 1} \sum_{j=1}^{\mathbb{N}_i} (\mathcal{F}^{(j)}(x^{(i)}) - \bar{\mathcal{F}}(x^{(i)}))(\mathcal{F}^{(j)}(x^{(i)}) - \bar{\mathcal{F}}(x^{(i)}))^T \\ &= M(V_{\mathcal{F}}(x)) \oplus T(x). \end{aligned} \tag{1.6.5}$$

The analysis then proceeds in two stages, which are individually similar to the deterministic simulator case, except there is a aleatory noise component, so the uncertainty in the simulator output never goes to zero. The first stage updates the population variance and the second stage updates the population mean. For the uncertainty in the unbiased estimator of the population mean at a location, the adjusted expectation of the population variance is used as a point estimate. Given $\mathcal{F}$ has a scalar output, the second order

specification for $D^{\bar{\mathcal{F}}}$ is approximated by

$$\mathbb{E}[\bar{\mathcal{F}}(x)] = \mathbb{E}[M(\mathcal{F}(x))], \tag{1.6.6}$$

$$\mathbb{E}[\hat{\Sigma}_{\mathcal{F}}(x)] = \mathbb{E}[M(V_{\mathcal{F}}(x))], \tag{1.6.7}$$

$$\mathbb{V}ar[\bar{\mathcal{F}}(x)] \approx \mathbb{V}ar[M(\mathcal{F}(x))] + \frac{1}{\mathbb{N}_i}\mathbb{E}_{\hat{\Sigma}_{\mathcal{F}}(X)}[M(V_{\mathcal{F}}(x))], \tag{1.6.8}$$

$$\mathbb{V}ar[\hat{\Sigma}_{\mathcal{F}}(x)] = \mathbb{V}ar[M(V_{\mathcal{F}}(x))] + V_T(x), \tag{1.6.9}$$

$$\mathbb{C}ov[\bar{\mathcal{F}}(x), \bar{\mathcal{F}}(x')] = \mathbb{C}ov[M(\mathcal{F}(x)), M(\mathcal{F}(x'))], \tag{1.6.10}$$

$$\mathbb{C}ov[\hat{\Sigma}_{\mathcal{F}}(x'), \hat{\Sigma}_{\mathcal{F}}(x')] = \mathbb{C}ov[M(V_{\mathcal{F}}(x)), M(V_{\mathcal{F}}(x'))], \tag{1.6.11}$$

$$\mathbb{C}ov[M(\mathcal{F}(x)), \bar{\mathcal{F}}(x')] = \mathbb{C}ov[M(\mathcal{F}(x)), M(\mathcal{F}(x'))] \text{ and} \tag{1.6.12}$$

$$\mathbb{C}ov[M(V_{\mathcal{F}}(x)), \hat{\Sigma}_{\mathcal{F}}(x')] = \mathbb{C}ov[M(V_{\mathcal{F}}(x)), M(V_{\mathcal{F}}(x'))] \tag{1.6.13}$$

where $V_T(x)$ is the uncertainty in the unbiased estimator for the variance. As the adjustment is performed in two stages rather than a joint update, the second order specification is approximated. For details on how to specify beliefs about $V_T(x)$ see [28] chapter 8. Emulator prediction is approximated by by

$$\mathbb{E}_{D^{\bar{\mathcal{F}}}}[\mathcal{F}(x')] \approx \mathbb{E}_{\bar{\mathcal{F}}(X)}[M(\mathcal{F}(x'))] \tag{1.6.14}$$

$$\mathbb{V}ar_{D^{\bar{\mathcal{F}}}}[\mathcal{F}(x')] \approx \mathbb{V}ar_{\bar{\mathcal{F}}(X)}[M(\mathcal{F}(x'))] + \mathbb{E}_{\hat{\Sigma}_{\mathcal{F}}(X)}[M(V_{\mathcal{F}}(x')]. \tag{1.6.15}$$

Toy Example: Stochastic Emulator

To illustrate emulation of a stochastic simulator, the previous toy example is supplemented with a population variance component. The toy example is now defined by

$$M(\mathcal{F}(x)) = 0.1 + 0.5x + 0.1\sin(10x) \text{ and} \tag{1.6.16}$$

$$M(V_{\mathcal{F}(x)}) = 0.001 + 0.01x^2. \tag{1.6.17}$$

Emulators were built and analysed using the one and two stage approach. The training set consisted of the simulator evaluated at 30 inputs with 10 repetitions; a total of 300 simulator evaluations. The specification of the simulator's population mean was the same as the deterministic toy example. The specification of the simulator's population variance was the same as the simulator's population mean. The regression parameters and variance multipliers were fixed to their GLS estimates, key formula 2.2.1. For the one-stage approach, the local noise at every simulator evaluation was fixed at 0.0001. For the two stage approach, $V_T(x) = 0.000005$. Finally, once the emulators had been built, 1000 additional

samples from the stochastic emulator were made and the standardised errors(equation 2.1.1) were calculated.

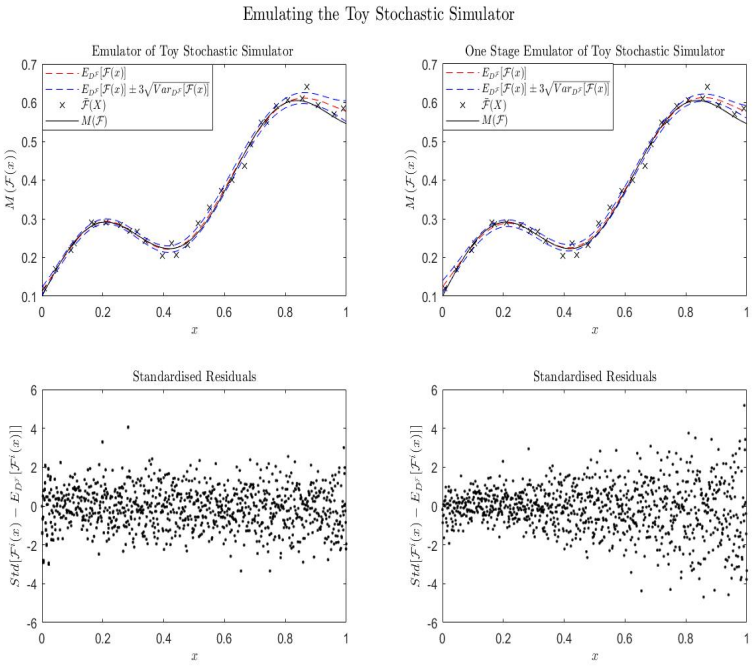

Figure A.6: Top left: adjusted beliefs for the toy stochastic simulator using the two stage update. Bottom left: standardised residuals for validation points using a two stage update. Top right: adjusted beliefs for toy stochastic simulator using a single stage update. Bottom right: standardised residuals for validation points using a single stage update.

Figure A.6 shows the difference between performing a two stage approach and updating the emulator in one stage with a constant noise term. The two stage approach better reflects uncertainty removing the *trumpet* shape in the standardised residuals for the validation points. Overall this method of emulating stochastic simulators is superior to ignoring the heteroscedasticity and has a good pain-to-gain ratio.

A.7 Bayes Linear Noisy Input-output Fields

In this section, a method for adjusting beliefs for a transfer function when the simulator output is uncertain will be presented. This method was used in early versions of the tuning algorithm presented in section 6.4.4. This is the better method when simulator uncertainty is significant compared to other sources of uncertainty. For simplicity, in this section, the simulator is assumed to have a scalar output.

When tuning a simulator, the difference between the actual simulator and its reified counterpart is described by $\Delta(\mathcal{F}(x))$. Given that $\mathcal{F}$ may be uncertain, a method for adjusting beliefs about $\Delta(\mathcal{F}(x))$ taking account of uncertainty in $\mathcal{F}(x)$ is required. The method presented in this section is analogous to [25] where Gaussian process regression models are adjusted with noisy input-output data.

It is assumed that the observations and the adjusted beliefs for the actual simulator may be written

$$z^{(i)} = \mathcal{F}^\star(x_S^{\star(i)}, x_E^\star, w^\star) \oplus \epsilon_{md}^{\star(i)} \oplus \epsilon_{me}^{(i)} \text{ and} \tag{1.7.1}$$

$$\mathbb{E}_{D^\mathcal{F}}[\mathcal{F}(x_S^{\star(i)}, x_E')] = \mathcal{F}(x_S^{\star(i)}, x_E') - \mathbb{A}_{D^\mathcal{F}}[\mathcal{F}(x_S^{\star(i)}, x_E')]. \tag{1.7.2}$$

Both $z^{(i)}$ and $\mathbb{E}_{D^\mathcal{F}}[\mathcal{F}(x_S^{\star(i)}, x_E')]$ have been substantiated. The relationship between the reified simulator and tuned simulator is specified as

$$\mathcal{F}^\star(x_S, x_E^\star, w^\star) = \Delta(\mathcal{F}(x_S, x_E')). \tag{1.7.3}$$

If the relationship is specified as $A = B + \Delta(B)$, it may always be reformulated into a form equivalent to equation 1.7.3. For example, if

$$A = \Delta(B) = B + \beta_0 + \beta_1 B + \mathcal{W}(B), \text{ then} \tag{1.7.4}$$

$$A = \Delta'(B) = \beta_0 + (\beta_1 + 1)B + \mathcal{W}(B). \tag{1.7.5}$$

Equation 1.7.3 implies that an observation may be written in the form

$$z^{(i)} = \Delta(\mathbb{E}_{D^\mathcal{F}}[\mathcal{F}(x_S^{\star(i)}, x_E')] + \mathbb{A}_{D^\mathcal{F}}[\mathcal{F}(x_S^{\star(i)}, x_E')]) \oplus \epsilon_{md}^{\star(i)} \oplus \epsilon_{me}^{(i)}. \tag{1.7.6}$$

Using a Taylor expansion, the observation may be approximated by

$$z^{(i)} \approx \Delta(\mathbb{E}_{D^\mathcal{F}}[\mathcal{F}(x_S^{\star(i)}, x_E')]) + \partial\Delta(\mathbb{E}_{D^\mathcal{F}}[\mathcal{F}(x_S^{\star(i)}, x_E')])\mathbb{A}_{D^\mathcal{F}}[\mathcal{F}(x_S^{\star(i)}, x_E')] \oplus \epsilon_{md}^{\star(i)} \oplus \epsilon_{me}^{(i)} \tag{1.7.7}$$

where $\partial\Delta(\mathbb{E}_{D^\mathcal{F}}[\mathcal{F}(x_S^{\star(i)}, x_E')])$ is the derivative of the transfer function evaluated at $\mathbb{E}_{D^\mathcal{F}}[\mathcal{F}(x_S^{\star(i)}, x_E')]$. The expectation and variance of z is then approximated by

$$\mathbb{E}[z^{(i)}] \approx \mathbb{E}[\Delta(\mathbb{E}_{D^\mathcal{F}}[\mathcal{F}(x_S^{\star(i)}, x_E')])], \tag{1.7.8}$$

$$\begin{aligned}\mathbb{V}ar[z^{(i)}] \approx\ & \mathbb{V}ar[\Delta(\mathbb{E}_{D^\mathcal{F}}[\mathcal{F}(x_S^{\star(i)}, x_E')])] \\ & + \partial\Delta(\mathbb{E}_{D^\mathcal{F}}[\mathcal{F}(x_S^{\star(i)}, x_E')])\mathbb{V}ar[\mathbb{A}_{D^\mathcal{F}}[\mathcal{F}(x_S^{\star(i)}, x_E')]]\partial\Delta(\mathbb{E}_{D^\mathcal{F}}[\mathcal{F}(x_S^{\star(i)}, x_E')])^T \\ & + \mathbb{V}ar[\epsilon_{me}^{(i)}] + \mathbb{V}ar[\epsilon_{md}^{\star(i)}] \text{ and}\end{aligned}$$

$$\mathbb{C}ov[z^{(i)}, z^{(j)}] \approx \mathbb{C}ov[\Delta(\mathbb{E}_{D^\mathcal{F}}[\mathcal{F}(x_S^{\star(i)}, x_E')]), \Delta(\mathbb{E}_{D^\mathcal{F}}[\mathcal{F}(x_S^{\star(j)}, x_E')]) + \mathbb{C}ov[\epsilon_{me}^{(i)}, \epsilon_{me}^{(j)}]. \tag{1.7.9}$$

To estimate the variance of $D^f = (z^{(1)}, \cdots, z^{(N)})$, a two stage approach is adopted. First the derivative is estimated assuming the adjusted version is zero. Second the adjusted expectation of the derivative is used as a point estimate to calculate the covariance matrix. For prediction at $\mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}(x_S^{\star\prime}, x_E')]$, the covariance is given by

$$\mathbb{C}ov[\Delta(\mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}(x_S^{\star\prime}, x_E')]), z^{(i)}] \approx \mathbb{C}ov[\Delta(\mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}(x_S^{\star\prime}, x_E')]), \Delta(\mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}(x_S^{\star(i)}, x_E')])]. \tag{1.7.10}$$

Then the adjusted beliefs for the reified simulator are calculated using

$$\mathbb{E}_{\{D^f, D^{\mathcal{F}}\}}[\mathcal{F}^\star(x_S^{\star'}, x_E^\star, w^\star)] \approx \mathbb{E}_{D^f}[\Delta(\mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}(x_S^{\star'}, x_E^{\mathcal{F}'})])] \text{ and} \tag{1.7.11}$$

$$\begin{aligned}\mathbb{V}ar_{\{D^f, D^{\mathcal{F}}\}}[\mathcal{F}^\star(x_S^{\star'}, x_E^\star, w^\star)] \approx{}& \mathbb{V}ar_{D^f}[\Delta(\mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}(x_S^{\star'}, x_E^{\mathcal{F}'})])] \\ &+ \mathbb{E}_{D^f}[\partial\Delta(\mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}(x_S^{\star'}, x_E^{\mathcal{F}'})])]\mathbb{V}ar_{D^{\mathcal{F}}}[\mathcal{F}(x_S^{\star'}, x_E^{\mathcal{F}'})] \\ &\mathbb{E}_{D^f}[\partial\Delta(\mathbb{E}_{D^{\mathcal{F}}}[\mathcal{F}(x_S^{\star'}, x_E^{\mathcal{F}'})])]^T.\end{aligned} \tag{1.7.12}$$

This techniques treats the adjusted version as essentially zero and adds a corrective term to the observation variance to compensate. Although there are approximations, this procedure for adjusting beliefs validates well and is efficient. $|\partial_x \Delta(x)|$ controls the amount of uncertainty, in the input, that is propagated through to uncertainty in the output. The higher this value the larger the propagated uncertainty in the output.

Derivative of Random Fields

In order to adjust beliefs for a noisy input-output model, it is necessary to adjust beliefs for the derivative of a random field. Here it is assumed that the field has a single variate and the indexing is a scalar, given by

$$y(x) = \sum_k \beta_{(k)} h_{(k)}(x) \oplus \mathcal{W}(x). \tag{1.7.13}$$

For scalar x, the derivative field $\partial_x y(x)$ is given by

$$\frac{\partial}{\partial x} y(x) = \partial_x y(x) = \sum_k \beta_{(k)} \partial_x h_{(k)}(x) \oplus \partial_x \mathcal{W}(x). \tag{1.7.14}$$

Beliefs for the derivative process at x' adjusted by $D^f = \{y(x_1), \cdots, y(x_N)\}$ are given by

$$\mathbb{E}_{D^f}[\partial_x y(x')] = \partial_x h(x') \mathbb{E}_{D^f}[\beta] + \mathbb{E}_{D^f}[\partial_x \mathcal{W}(x')] \text{ and} \tag{1.7.15}$$

$$\mathbb{V}ar_{D^{\mathcal{F}}}[\partial_x y(x')] = \partial_x h(x') \mathbb{V}ar_{D^f}[\beta] \partial_x h(x')^T + \mathbb{V}ar_{D^f}[\partial_x \mathcal{W}(x')] \tag{1.7.16}$$

where

$$\mathbb{E}_{D^{\mathcal{F}}}[\partial_x \mathcal{W}(x')] = \mathbb{C}ov[\partial_x y(x'), \mathcal{W}(X)] \mathbb{V}ar[D^{\mathcal{F}}]^{-1} (y - H\mathbb{E}[\beta]). \tag{1.7.17}$$

The covariance term is given by

$$\mathbb{C}ov[\partial_x y(x'), \mathcal{W}(X)] = \mathbb{C}ov[\partial_x \mathcal{W}(x'), \mathcal{W}(X)]. \tag{1.7.18}$$

For example, if

$$\mathbb{C}ov[\mathcal{W}(x), \mathcal{W}(x')] = \sigma \exp\left[-0.5\left(\frac{x-x'}{l}\right)^2\right] \tag{1.7.19}$$

then

$$\mathbb{C}ov[\mathcal{W}(x), \partial_x \mathcal{W}(x')] = -\sigma \frac{x-x'}{l^2} \exp\left[-0.5\left(\frac{x-x'}{l}\right)^2\right]. \tag{1.7.20}$$

The final calculation that is required is the prior uncertainty in the derivative process which is found by taking the derivative of the covariance function with respect to x and x'. For the covariance function in equation 1.7.19, this is given by

$$\mathbb{C}ov[\partial_x \mathcal{W}(x), \partial_x \mathcal{W}(x')] = \left(\frac{1}{l^2} - (x-x')^2 \frac{1}{l^4}\right) \mathbb{C}ov[\mathcal{W}(x), \mathcal{W}(x')]. \tag{1.7.21}$$

Then the adjusted variance of the residual process is given by

$$\begin{aligned} \mathbb{V}ar_{D^f}[\partial_x \mathcal{W}(x')] = \mathbb{C}ov[\partial_x \mathcal{W}(x'), \partial_x \mathcal{W}(x')] - \\ \mathbb{C}ov[\partial_x y(x'), \mathcal{W}(X)] \mathbb{V}ar[D^f]^{-1} \mathbb{C}ov[\partial_x y(x'), \mathcal{W}(X)]^T. \end{aligned} \tag{1.7.22}$$

Derivatives of random fields may be extended to multiple inputs and multiple outputs, see [52,116] or [54] for details. If the derivative of a random field is required, the choice of covariance specification is limited to those which result in differentiable random fields.

Toy Derivative of a Random Field

To illustrate a derivative process we return to the toy example used to illustrate a deterministic simulator. Figure A.7 shows adjusted beliefs for the toy simulator and its derivative. An important feature to note is, even if the uncertainty on the field itself is zero, the uncertainty in the derivative never goes to zero. The is because although a location on the field may be known, the gradient of the slope always requires some uncertain locations to be included in its calculation.

Toy Transfer Function

To illustrate a noisy input-output transfer function, the toy deterministic simulator is complemented with a independent noise process on both inputs and outputs. The model

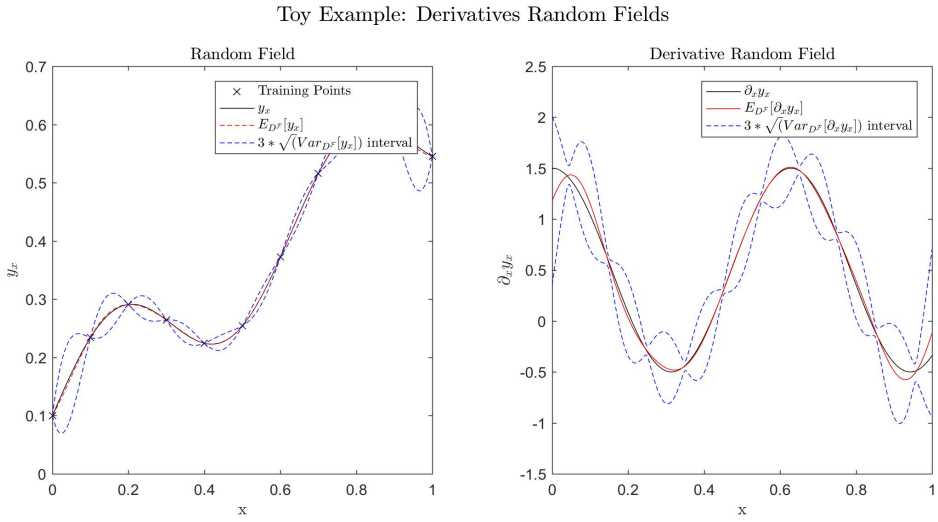

Figure A.7: Left: adjusted beliefs for toy example. Right: adjusted beliefs for serivative field.

is described by

$$\Delta(\mathcal{F}(x)) = 1 + 0.5\mathcal{F}(x) + 0.1\text{sin}(\mathcal{F}(x)) \text{ and} \tag{1.7.23}$$

$$z = \Delta(\mathcal{F}(x)) + \epsilon_{me} + \epsilon^{\star}_{md}. \tag{1.7.24}$$

To simulate uncertainty in $\mathcal{F}(x)$, a sample from a normal distribution was made with standard deviation 0.02 and added to the input. The measurement error and model discrepancy was sampled from a normal distribution with variance 0.01. Figure A.8 shows a comparison between adjusted beliefs, assuming that the inputs are certain and a noisy input-output adjustment. The top line shows the adjusted beliefs for the transfer function. The bottom line shows the standardised residuals for the adjusted version of the reified emulator. The adjusted variance is underestimated in the steep regions for the naive approach meaning that the known truth has a high standardised error. The advantage of including noisy input output models is subtle as the function inferred in both cases is similar. However, when the uncertainty estimation is considered the advantage is clear.

A.8 Dose Considerations

In chapter 6, it was shown that there is evidence of image quality improvement and inferences may be made about the entity composition. A major motivation for this work was to be able to reduce the dose a patients receives in a standard X-ray exposure. The dose a patient receives is usually decided by the signal-to-noise-ratio in the image. Signal-to-noise-ratio is preferred by radiologists to variance and they often refer to it as variance

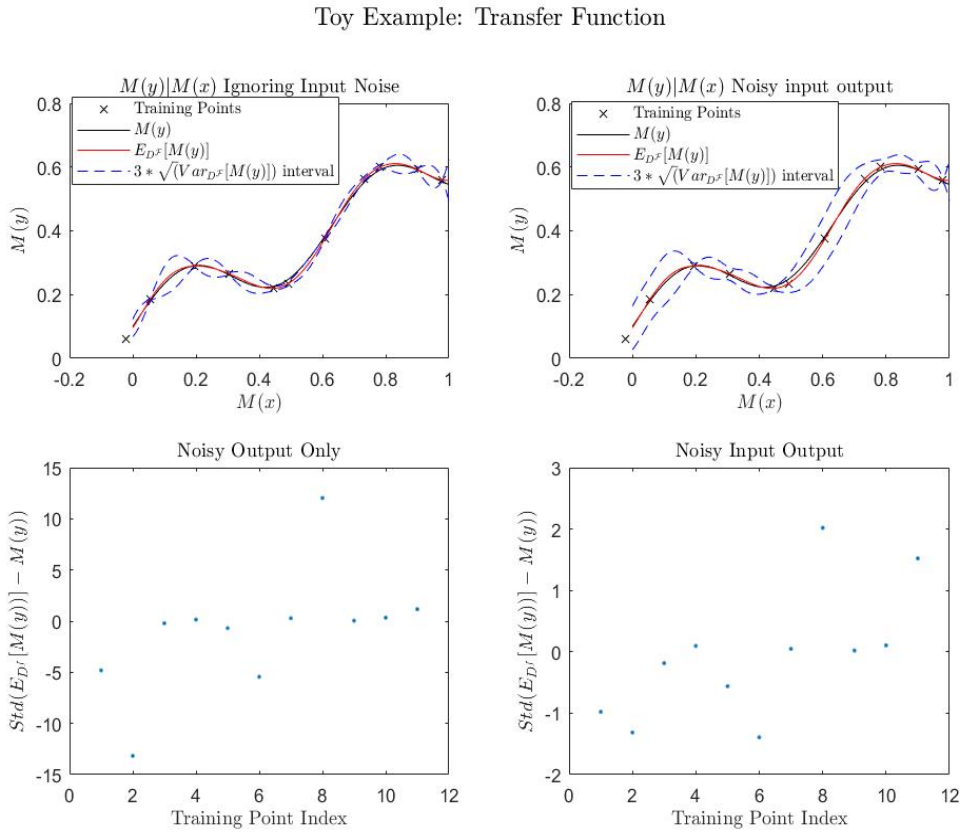

Figure A.8: Top left: adjusted beliefs assuming no input noise. Top right: adjusted beliefs assuming noisy inputs. Bottom left: standardised residuals assuming no input noise. Bottom right: standardised residuals assuming noisy inputs.

or noise. For example, increasing the dose decreases the variance or noise. This can be misleading as the variance of the local component of the noise increases as does the signal-to-noise ratio. Their preference is because the absolute variance is not of direct interest for image quality, only how features may be differentiated from local noise. Signal-to-noise ratio for a random variable X is given by

$$SNR(X) = \frac{\mathbb{E}[X]}{\sqrt{\mathbb{V}ar[X]}}. \tag{1.8.1}$$

Given the primitive process in X-ray imaging is assumed to be a counting process, the local noise component is assumed to be approximately Poisson. Therefore, the signal-to-noise ratio of an X-ray image is usually assumed to be

$$SNR(X) = \frac{\mathbb{E}[X]}{\sqrt{b\mathbb{E}[X]}} \tag{1.8.2}$$

where b is a constant which makes the variance proportional to the expectation rather than equalling it. If the dose is increased by factor c, the signal-to-noise ratio is given by

$$\frac{c\mathbb{E}[X]}{\sqrt{c}\sqrt{b\mathbb{E}[X]}} = \sqrt{c}SNR(X). \tag{1.8.3}$$

Overall this model is a good representation and is well supported by experimental data. In order to describe the model for dose comparison, the variance of the local noise of an X-ray image is split into direct and scattering components.

$$\mathbb{V}ar[X] = \mathbb{V}ar[X_D] + \mathbb{V}ar[X_S]. \tag{1.8.4}$$

Then the variance of an image with an ASG is assumed to be

$$\mathbb{V}ar[X_{ASG}] = d\mathbb{V}ar[X_D] + d'\mathbb{V}ar[X_S] + \mathbb{V}ar[X_{ASG-Art}]. \tag{1.8.5}$$

Here d and d' are the proportion of photons blocked by the grid for direct and scattering components respectively. $\mathbb{V}ar[X_{ASG-Art}]$ is the additional local noise which is added from ASG artefacts. The artefacts are not random noise but usually an aliasing pattern; they do however degrade image quality and will affect the variance calculated so it is logical to include it in the model. Then the signal-to-noise ratio is given by

$$SNR(X_{ASG}) = \frac{d\mathbb{E}[X_D] + d'\mathbb{E}[X_S]}{\sqrt{d\mathbb{V}ar[X_D] + d'\mathbb{V}ar[X_S] + \mathbb{V}ar[X_{ASG-Art}]}}. \tag{1.8.6}$$

To retain diagnostic image quality using an ASG, radiologists typically increase the dose by a factor c such that

$$SNR(X_{ASG}) = SNR(X). \tag{1.8.7}$$

For the scatter correction method proposed, only the low frequency components are effected and the local noise is not. Therefore,

$$SNR(X_{SC}) = \frac{\mathbb{E}[X_D] + \tilde{d}\mathbb{E}[X_S]}{\sqrt{\mathbb{V}ar[X_D] + \mathbb{V}ar[X_S]}} \tag{1.8.8}$$

where $\tilde{d}$ is the proportion of scatter remaining in the expectation. A dose benefit is obtained when

$$SNR(X_{SC}) > SNR(X_{ASG}) \tag{1.8.9}$$

for an equivalent patient dose. The simplest case of this analysis is when $\tilde{d} = d' = \mathbb{V}ar[X_{ASG-Art}] = 0$. Then equation 1.8.9 may be rearranged to give

$$\frac{1}{\sqrt{1 + \frac{\mathbb{V}ar[X_S]}{\mathbb{V}ar[X_D]}}} \frac{\mathbb{E}[X_D]}{\sqrt{\mathbb{V}ar[X_D]}} > \sqrt{d}\frac{\mathbb{E}[X_D]}{\sqrt{\mathbb{V}ar[X_D]}} \tag{1.8.10}$$

as $\frac{\mathbb{E}[X_D]}{\sqrt{\mathbb{V}ar[X_D]}}$ is strictly greater than zero. Then the critical point is

$$(1 + SPR) = d^{-1} \tag{1.8.11}$$

where $SPR = \frac{\mathbb{V}ar[X_S]}{\mathbb{V}ar[X_D]}$ is referred to as scatter-to-primary-ratio and d^{-1} is the amount of additional dose required to make an ASG have diagnostic quality SNR. So, for a grid which blocks 50% of the direct beam and for entities with a scatter-to-primary-ratio less than one, a dose benefit is obtained. The reason this case is used is it is simple enough to be understood and reduces the problem to two parameters, which are calculable. However, by simplifying the argument major features of analysis are lost and the model does not match experimental findings.

Here a simply adjustment to the conventional model is proposed. The dose improvement factor is calculated as

$$\mathcal{F}^{SNR}(SPR, d, d') = \left(\left(\frac{\mathbb{E}[X_D]}{\sqrt{\mathbb{V}ar[X_D] + \mathbb{V}ar[X_S]}} \right) \left(\frac{d\mathbb{E}[X_D]}{\sqrt{d\mathbb{V}ar[X_D] + d'\mathbb{V}ar[X_S]}} \right)^{-1} \right)^2 . \tag{1.8.12}$$

The transmission of scatter is assumed in both cases to be zero. In the ASG case, it would be expected that the transmitted scatter is greater than zero but this would penalise the scatter corrected case. As scatter remaining in the image would decrease the ability of clinicians to see small features, it seems inappropriate to penalise the scatter corrected case. This feature should penalise the ASG case and increasing dose would not fix this problem. Similarly, although $\mathbb{V}ar[X_{ASG-Art}]$ effects image quality and signal-to-noise ratio, it cannot be rectified by a simple dose increase. The focus of this section is dose, not general image quality, so factors which may not be improved by a simple dose increase are removed. Overall, the model is fairer but still favours the ASG as artefacts and the fact that $\mathbb{E}[X_D - \frac{1}{d}X_{ASG}] < 0$ is not considered.

Figure *A*.9 shows analysis of this dose model. The yellow regions are where there is a dose benefit for scatter correction over ASG and the non-yellow regions indicate a dose deficit. This analysis poses interesting questions about how dose should be chosen for patients based on signal-to-noise arguments. The factor by which the dose is increased is essentially based only on attaining the same level of intensity as the no scatter case. However, it is questionable whether this is worthwhile when two main sources of error present are not reduced by increasing the dose. Suppose that an object has scatter to primary of 4 and the grid blocks 50% of direct photons and 87.5% of scattered photons. Then the resulting ASG image has a scatter-to-primary-ratio of 1. As increasing the dose

for an image with scatter-to-primary-ratio of 1 is not considered an improvement, it is hard to argue why it is in this case.

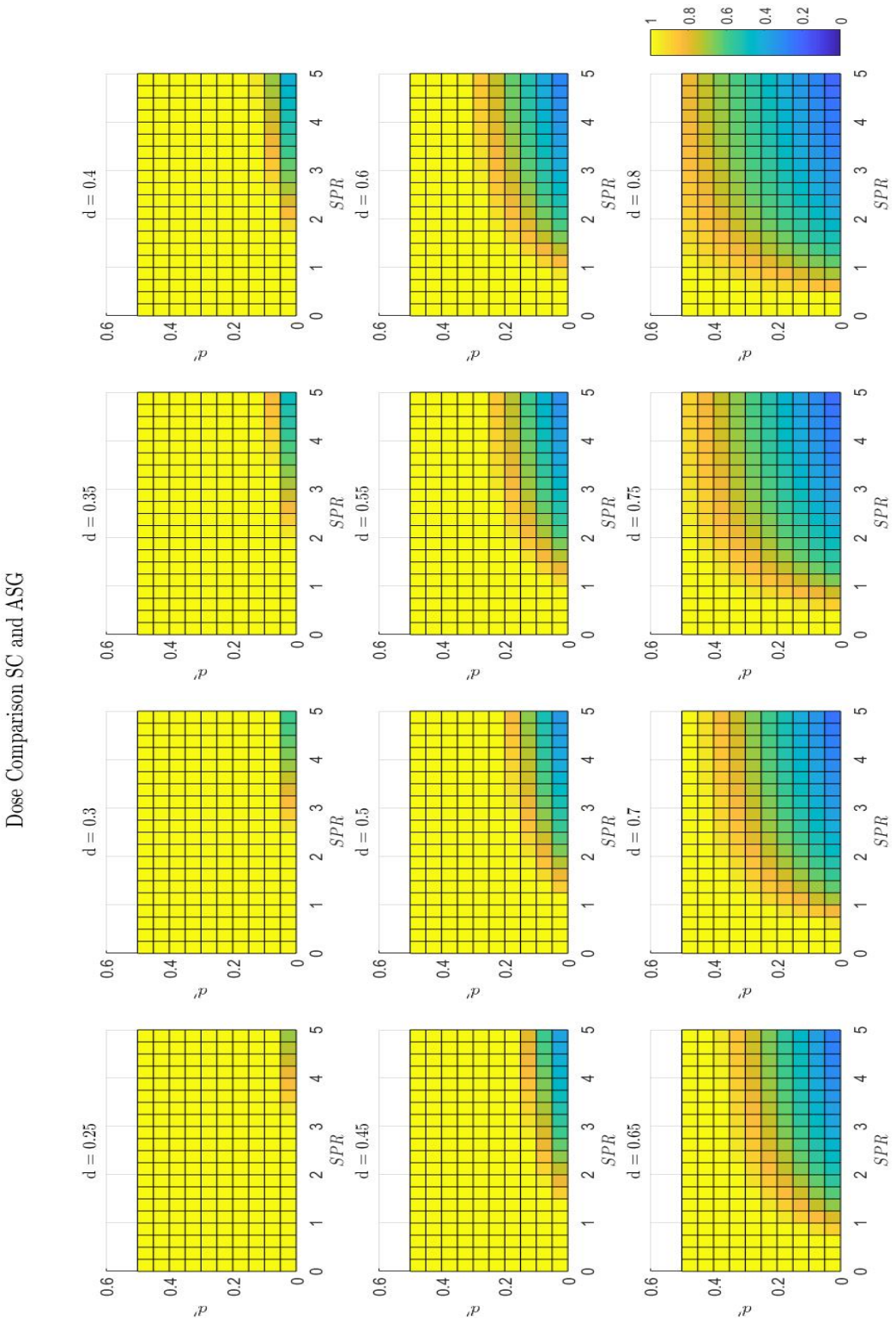

Figure A.9: Heat map of does impact for varying entity dependent parameters.

A.9 Structural Tuning of GEANT4

In this section, an illustrative example of structural tuning will be presented. This is intended to exemplify the techniques and not as a complete analysis. As such, we make two simplifications. This first is that all analysis is done on the central pixel of the detector. A full analysis would use the entire output image. The second is that only improvements to the simulator, for which the simulator itself could be evaluated, were considered. The process will show that the variation induced by kV and filtration cannot be tuned out but variable scintillator thickness and changes in filtration material can. Then second order specifications will be made for a diagnostic quantity using internal analysis. Finally, actual non-implausible sets of tuning parameters and diagnostics are calculated using real world data in an external analysis. Finally, comparisons are made between the internal and external analyses.

The transfer function assumptions and specifications were assumed to be the same as in section 6.4. The sources of discrepancy are from parameters which have been fixed to their *a-priori* expected system values when they should be uncertain. The process proposes an improved simulator by releasing an equipment dependent parameter. For each proposed improved simulator, an emulator is built using evaluations of the simulator, or the simulator itself is evaluated. That emulator is then used to produce a set of calibration training-sets. For every calibration training-set, the non-implausible set of tuning parameters and transfer functions is calculated using the method described in section 6.4. See algorithm 15 for further details.

Adding kV

To illustrate the technique, it is assumed that the current version of the simulator available is

$$\mathcal{F}(x_S) \tag{1.9.1}$$

that is the simulator from section 6.4 with kV and filtration at the source fixed to their *a-priori* expected values. The first improved simulator considered is defined as

$$\mathcal{F}'(x_S, kV) \tag{1.9.2}$$

where $\mathcal{F}'(x_S, 80) = \mathcal{F}(x_S)$. In other words, we are checking whether the variation that could be induced by the uncertainty in kV can be resolved by the transfer function. The prior bounds for kV are set as $[77, 83]$. A sample of points are made from this interval

and a calibration training-set emulated for every kV sampled. In this case, the calibration training-set was 10 cuboid entities of aluminium and PMMA with thicknesses $[1, 2, 3, 4, 5]$ and $[5, 10, 15, 20, 25]$.

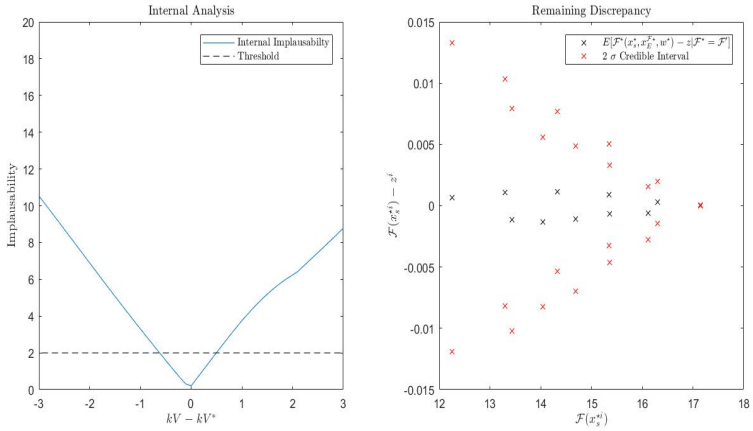

Figure A.10: Left: implausibility for tuning $\mathcal{F}$ to $\mathcal{F}'$ for calibration training-sets at a number of kVs. Right: expected and uncertainty bounds for the functional discrepancy between $\mathcal{F}$ and $\mathcal{F}^{\star}$.

Figure A.10 shows two outputs from the internal analysis of the improved simulator. The left hand plot shows the implausibility when tuning $\mathcal{F}$ to $\mathbb{E}[\mathcal{F}']$. With an implausibility threshold of 2, there is a large portion of the prior non-implausible space that $\mathcal{F}$ cannot be tuned to $\mathbb{E}[\mathcal{F}']$. This can be interpreted as, if $\mathcal{F}^{\star} = \mathbb{E}[\mathcal{F}']$, the uncertainty in kV must be reduced to $[kV^{\star} - 0.5, kV^{\star} + 0.5]$ to expect to be able to tune $\mathcal{F}$ to $\mathcal{F}'$. The left hand of the plot shows the expected and variance in the remaining discrepancy for points within the region $[kV^{\star} - 0.5, kV^{\star} + 0.5]$. This shows that, if i) $\mathcal{F}^{\star} = \mathbb{E}[\mathcal{F}']$ and ii) the prior uncertainty can be reduced to $[kV^{\star} - 0.5, kV^{\star} + 0.5]$ – what we expect the remaining discrepancy to be. Given that within the prior non-implausible set there are many points for which it is not expected that a tuning input can be found, further improvements are required to the simulator before it is worth spending resource on making observations. This is confirmed by attempting analysis with real world observations and getting an implausibility of 10.38.

Adding Filtration

As $\mathcal{F}$ cannot be tuned to $\mathcal{F}'$, the current simulator is now considered $\mathcal{F}'$ and the next improvement is to allow filtration to become variable. In words, we are checking whether making kV a tuning parameter means that the discrepancy induced by fixing filtration, to a value which might not be its system value, can be resolved. The second improved

simulator is given by

$$\mathcal{F}''(x_S, kV, fil) \tag{1.9.3}$$

where $\mathcal{F}''(x_S, kV, 2) = \mathcal{F}'(x_S, kV)$. To build an emulator of $\mathcal{F}''$, the simulator was evaluated at a number of kV and filtrations for the calibration training entities.

Given the actual filtration is from many different materials, the effective thickness of aluminium filtration is likely to differ from the reported values but not by a large amount. For illustrative purposes, the prior non-implausible set for filtration at the source is specified as $[0, 4]$. Then for a range of kV and filtrations, the non-implausible sets of the tuning parameters are calculated for tuning $\mathcal{F}'$ to $\mathbb{E}[\mathcal{F}'']$. Figure A.11 illustrates

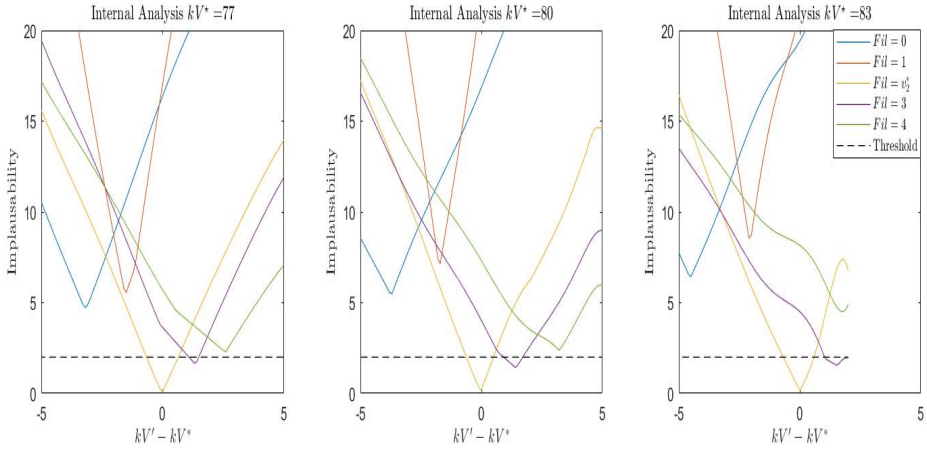

Figure A.11: Left: implausibility profiles for tuning $\mathcal{F}'$ to $\mathcal{F}''$ when $kV^\star = 77$. Right: implausibility profiles for tuning $\mathcal{F}'$ to $\mathcal{F}''$ when $kV^\star = 83$. Centre: implausibility profiles for tuning $\mathcal{F}'$ to $\mathcal{F}''$ when $kV^\star = 80$.

some features of the output of tuning $\mathcal{F}'$ to $\mathbb{E}[\mathcal{F}'']$. There is large portions of the prior non-implausible space of filtration where $\mathcal{F}'$ cannot be tuned to $\mathbb{E}[\mathcal{F}'']$. From this we can infer that, given filtration could be anywhere in $[0, 4]$, there are large portions of $[0, 4]$ where $\mathcal{F}'$ cannot be tuned to $\mathcal{F}^\star$ and therefore it is unlikely that $\mathcal{F}'$ can be tuned to $\mathcal{F}^\star$. Figure A.11 also shows that changing kV is better able to take the discrepancy into account when $fil^\star$ is greater than the value it has been fixed to in $\mathcal{F}'$. This was not something which we *a-priori* thought was true but something we learnt from the analysis. From performing this analysis at higher resolution in fil, the prior non-implausible set for the tuning parameters could be chosen. This was given by the range $[76, 85]$ and calculated by finding the largest and smallest kV's in all non-empty posterior non-implausible sets.

Figure A.12 shows the analysis if it was decided to tune $\mathcal{F}'$ to $\mathcal{F}^\star$. There are no

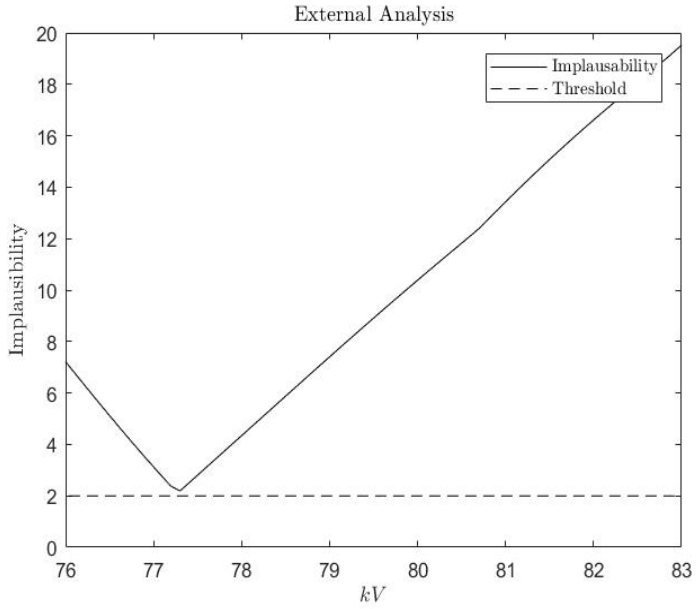

Figure A.12: implausibility vee profiles tuning $\mathcal{F}'$ to $\mathcal{F}^\star$ using real world observations.

non-implausible points but the min-implausible point is within the $[76, 85]$ region. It is tempting to state that this implies the system value for filtration is close to 2. However, given that $\mathcal{F}^\star \neq \mathcal{F}''$, this would be misleading as other types of discrepancy between $\mathcal{F}^\star$ and $\mathcal{F}''$ could be compensated for by fixing filtration to the wrong system value.

Adding Scintillator Thickness

The next source of discrepancy that will be considered is scintillator thickness. Manufactures report a scintillator thickness but it is difficult to control the production process exactly. Therefore, the next improved simulator considered is given by

$$\mathcal{F}'''(\eta, \mu, kV, fil, Sc) \tag{1.9.4}$$

where

$$\mathcal{F}'''(\eta, \mu, kV, fil, 500) = \mathcal{F}''(\eta, \mu, kV, fil). \tag{1.9.5}$$

In words, we are trying to assess whether allowing kV and filtration to vary as tuning parameters, enables the resolution of discrepancy induced by fixing scintillator thickness to a value which may not be its system value.

The prior non-implausible set for Sc was $[400, 600]\mu m$ and the prior expected value was $Sc^\star = 500\mu m$ as reported by the manufacturer. For this demonstration, instead of varying $kV^\star$ and $fil^\star$ they were fixed to their *a-priori* system values. Figure A.13 shows

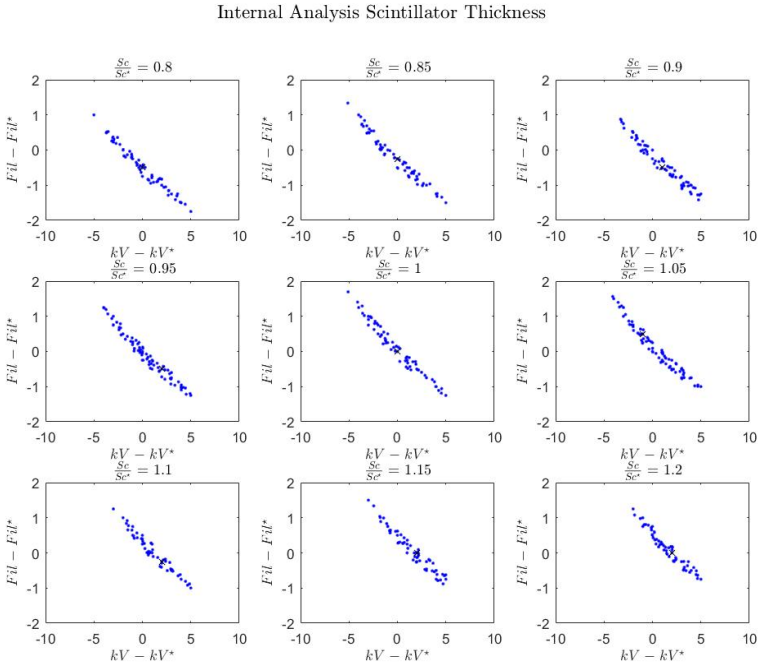

Figure A.13: Samples from posterior non-implausible sets of tuning parameters for a range of scintillator thicknesses.

the non-implausible sets, for tuning $\mathcal{F}''$ to $\mathbb{E}[\mathcal{F}''']$, for a variety of scintillator thicknesses from the prior non-implausible set. Every tested point has a non-empty posterior non-implausible set. From this we can infer that the discrepancy induced by not including scintillator thickness may be tuned out.

Figure A.14 shows the expectation and approximate credible interval of the difference between the summary statistic of the tuning input and system values. The summary statistic in this case was the mean of the non-implausible set. It appears to favour higher kV and lower filtration and there is negative correlation between the two.

Figure A.15 illustrates a second order specification for the scintillator thickness and tuning parameters. There is a clear positive correlation between increasing scintillator thickness and an increase in filtration. This is mildly unexpected as increasing scintillator thickness should have the opposite effect on the output as increasing filtration but, given the complex functional relationship and the transfer function, it is not entirely implausible. There is correlation between the scintillation thickness and tuning parameters so, even if the Sc is not included as an input to the simulator, beliefs about it may be incorporated by adjusting beliefs for the centre of the non-implausible set.

Figure A.16 illustrates how the outputs of the internal analysis may be utilised to

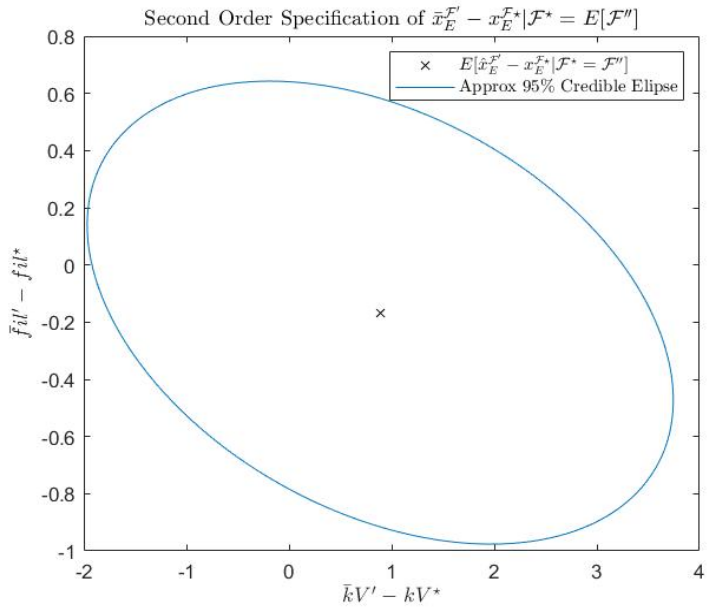

Figure A.14: : Illustration of second order specification of beliefs about a summary statistic from the non-implausible set.

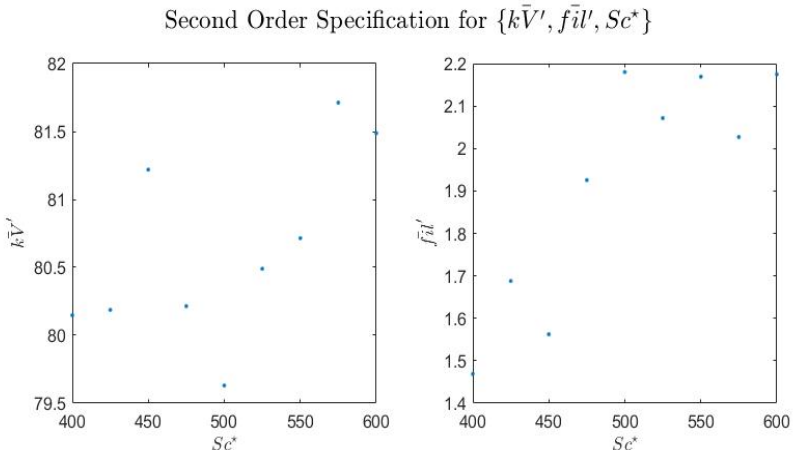

Figure A.15: Plot showing the relationship between varying scintillator thickness and the mean of the non-implausible set.

specify a prior non-implausible set for the tuning parameters. The plot shows the points from every non-implausible set and their convex hull. Therefore, given the system values of $kV^\star$ and $fil^\star$ are $[80, 2]$, to resolve the discrepancy from not including scintillator thickness in the simulator, the prior non-implausible set for tuning parameters must be at least as big as the convex hull in figure A.16.

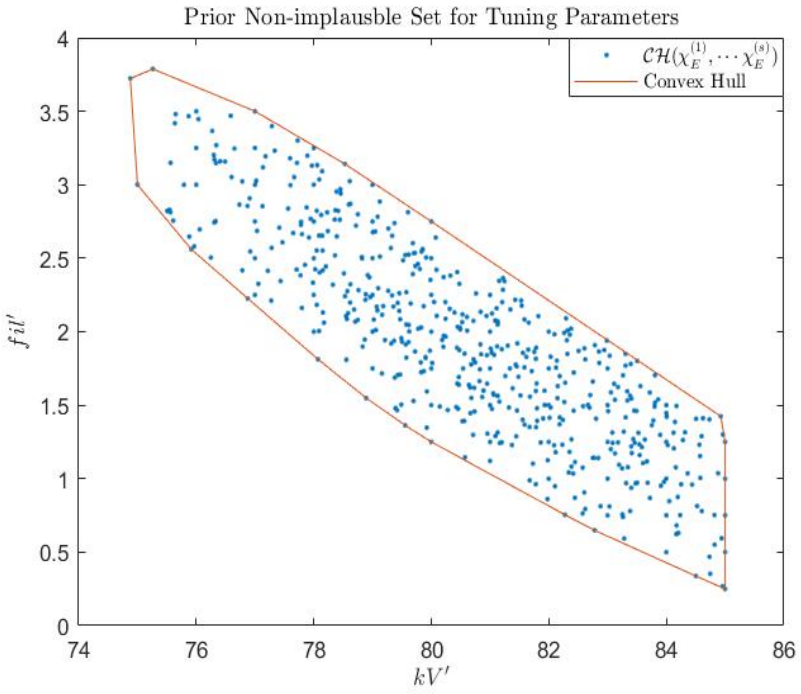

Figure A.16: Prior non-implausible set for tuning parameters given $\mathcal{F}^\star = \mathbb{E}[\mathcal{F}''']$

Adding Copper Filtration

The next improved simulator will address the material of the filter. The filtration at the source given by manufacturers is effective thickness of aluminium. However, in reality there will be a number of materials at the source. Therefore, an improved simulator would have more than one material in the filter. This is added to the simulator by adding a copper sheet just after the aluminium filter. The improved simulator is given by

$$\mathcal{F}''''(\eta, \mu, kV, fil, Cu\eta) \tag{1.9.6}$$

where $\mathcal{F}''''(\eta, \mu, kV, fil, 0) = \mathcal{F}''(\eta, \mu, kV, fil)$. Once again for the demonstration the system values of $kV^\star$ and $fil^\star$ were fixed to their *a-priori* system values. In words, given that $kV^\star$ and $fil^\star$ are $[80, 2]$, can making kV and fil tuning parameters, resolve the discrepancy induced by fixing $Cu\eta$ to a value which may not be the system value. To create the calibration training-sets for the internal analysis, the simulator was evaluated for a number of copper filtrations creating a set of simulated calibration training sets.

Figure $A.17$ shows the non-implausible sets for a range of additional copper filtration. As copper is around 3.33 times as dense as aluminium, it was expected that 0.16mm of copper would be counteracted by around 0.53mm of aluminium. In fact, the amount of additional aluminium required is much larger. For 0.1mm of copper the edge of the non-

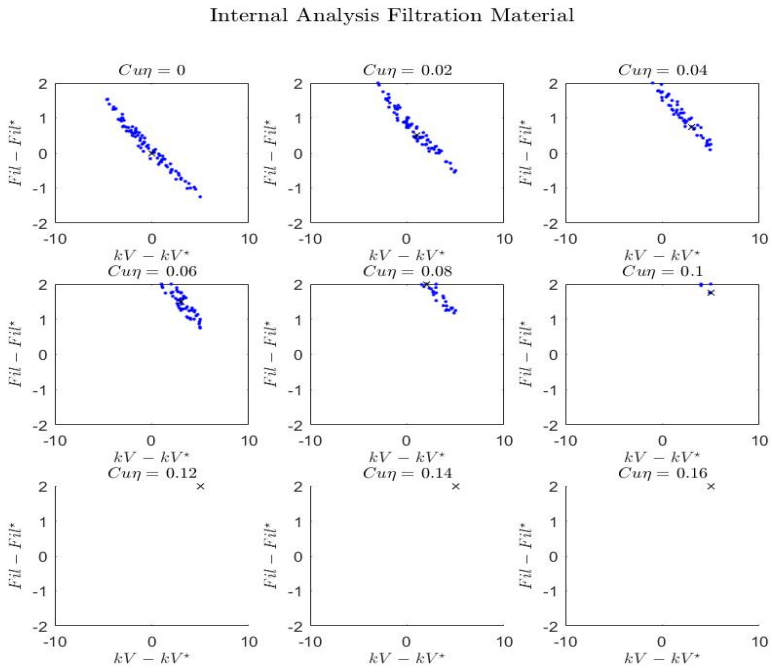

Figure A.17: Non-implausible set of tuning parameters for a set of additional copper filtrations.

implausible filtration is two away from the system value and the kV is 5 away from its system value. Therefore, if there is copper filtration in the beam, the expected aluminium thickness is higher than the relative density difference. Furthermore, if the material is variable the prior non-implausible set of fil' should be much larger than $[0, 4]$.

External Diagnostics

Given the internal analysis, it is expected that allowing kV and fil to become tuning parameters will enable $\mathcal{F}''$ to be tuned to $\mathcal{F}''''$. Given that any subsequent improvements are expected to be smaller than those considered so far, we expect to be able to tune $\mathcal{F}''$ to $\mathcal{F}^\star$. An external analysis will now be performed with real world data and comparisons made between internal and external analysis.

Figure A.18 shows the total set of non-implausible points, given the system values for kV is 80 and filtration is 2, for all improved simulators and their convex hull. The real world non-implausible points are contained within the convex hull except at the bottom right corner. This indicates that the actual non-implausible set of tuning parameters is not in conflict with the prior specification using improved simulators.

Figure A.19 shows a comparison of expected and revealed summary statistics from

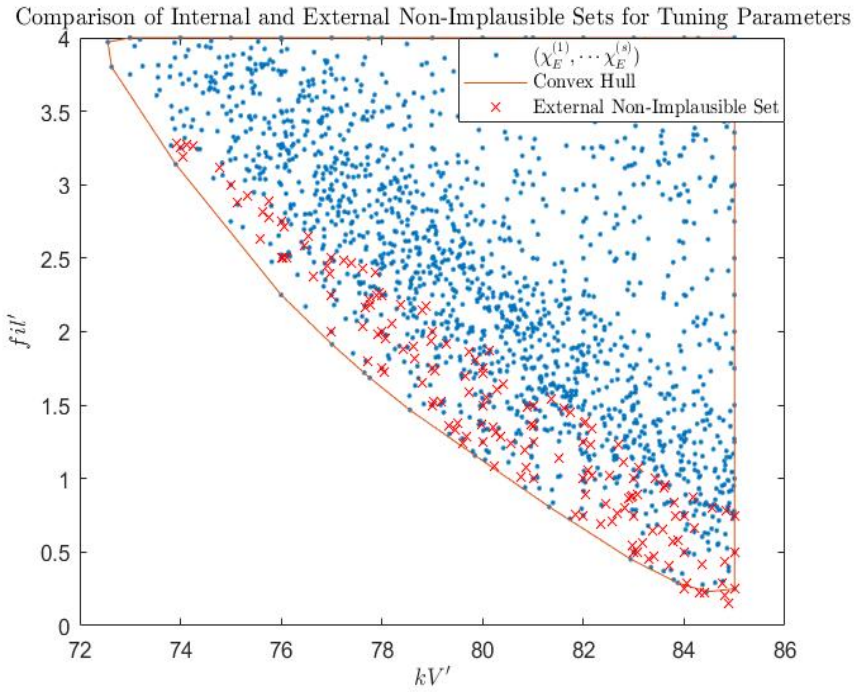

Figure A.18: Comparison of the internal and external non-implausible sets.

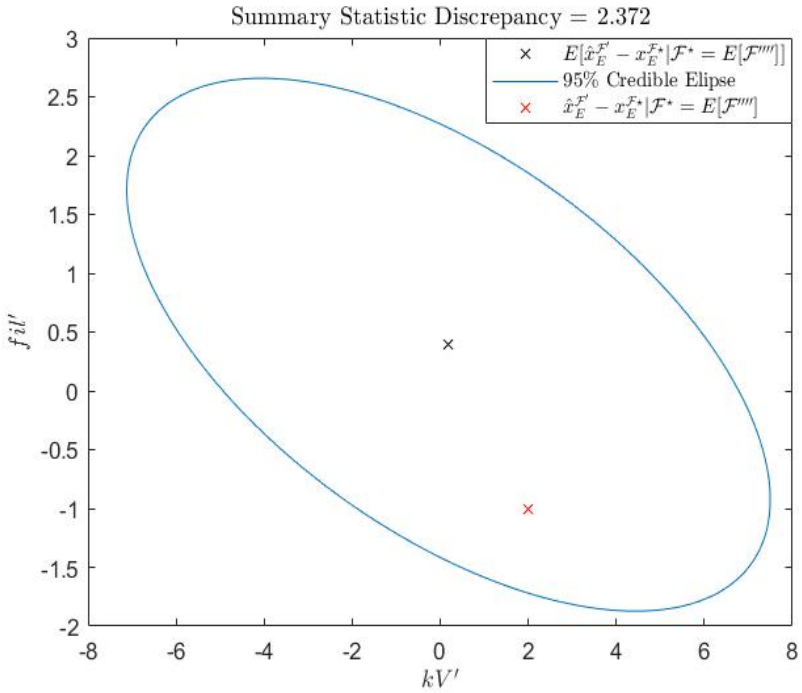

Figure A.19: Comparison of second order specification of a diagnostic quantity and the actual diagnostic calculated with real world data.

the non-implausible sets. Again there is no evidence to suggest that there is a further significant source of discrepancy which is invalidating the tuning process. The result of structural tuning is that it is appropriate to tune $\mathcal{F}''$ to the reified counterpart and the tuned version is applicable to the real world. Furthermore, there is a toolbox of techniques to explore different sources of discrepancy and assess whether they are likely to have a high impact on the tuning approach. If a non-implausible set is not found, this provides further information as to the source of problem. For example, if the centre of the posterior non-implausible set for tuning parameters is higher than expected, it indicates that another material, possibly copper, is used for filtration at the source. This was helpful when calibrating to experimental systems we were less familiar with. In one scenario, the min-implausible point had a large filtration, we then asked the manufacturer whether there was any other filtration in front of the source and they confirmed that copper had been added.

This section has exemplified structural tuning by assessing a few of the discrepancies present in our build of G4. The method provides a framework for assessing whether a simulator can be tuned and specifying beliefs for the quantities involved.