
GYROID BEAMS: A STOCHASTIC ANALYSIS

A PREPRINT

Ben Boys
Gonville & Caius College, Cambridge
bb515@cam.ac.uk

September 1, 2019

1 Introduction

Stochastic modelling has been applied to Engineering designs to assess how they are behaving in the real world and to provide feedback for new design iterations. For example, the world's first single-span, 3D printed steel bridge is fitted with sensors to measure strain, acceleration, vibrations and temperature, with new data generated whenever the bridge is traversed [1]. This report will apply a stochastic analysis to a well known problem: deflection of a simply supported beam. It aims to show how applying stochastic methods to a known Engineering problem can generate Bayesian parameters that well explain the data.

The structure that will be modelled is a shell structure. Thin shells are curved solids with one dimension significantly smaller than the other two. They are prevalent in nature, e.g., as insect wings or biological membranes and in engineering, most prominently in aerospace and automotive. Carefully designed curved thin shells have a load carrying capacity which is usually significantly higher than comparable flat structures.

We consider the composite beam shown in Figure 1 consisting of a gyroid core and two face plates. The gyroid is a triply periodic minimal surface with zero mean curvature and has recently been extensively explored in additive manufacturing applications, see e.g. Hussein et al. (2013) [2]; Abueidda et al. (2017) [3]. As known, cellular solids like the gyroid core can have mechanical properties that are orders of magnitude different from their constituent materials (Fleck et al., 2010) [4]. The length of the beam is 0.243m, its height, i.e. distance between the top and bottom plates, is 0.031m and its width is 0.031m. The gyroid core is described by the algebraic function

$$\sin(\lambda x)\cos(\lambda y) - \sin(\lambda y)\cos(\lambda z) - \sin(\lambda z)\cos(\lambda x) = 0 \quad (1)$$

with $\lambda = 20\pi$. The core and the two plates are modelled as thin shells and have a thickness of $t = 0.906$ mm.

The beam is simply supported at a distance 11mm away from the boundaries (giving a span of 220mm). The top plate is subjected to a line load of total distributed force W acting in the negative z direction at mid span.

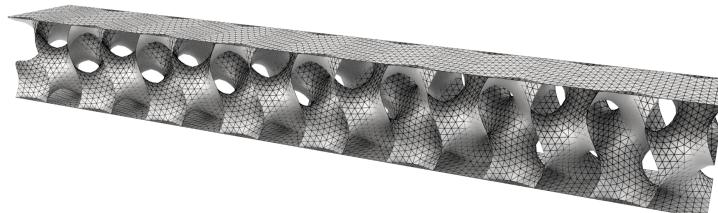


Figure 1: The composite, thin shelled beam that will be studied in this report, which consists of a gyroid core sandwiched between two face plates.

The next section will present two different models that describe the deflection of the beam under in plane loading: modelling it as a thin shell, using the Finite Element Method (FEM) and as an I-beam using Euler-Bernoulli beam theory. The beam was 3D printed so that deflection data could be collected: Section 3 describes this process. Section 4 will demonstrate a stochastic approach to assess quantitatively how well the two different beam models predict the deflection of the 3D printed beam. Once the best ranking model has been chosen, Section 5 will stochastically model the mismatch error between the models own predicted deflections and the observed deflections, employing Monte Carlo procedures. Section 6 presents a discussion of the results and future work.

2 Beam Models

2.1 Euler-Bernoulli Model

The deflection of the gyroid structure can be approximated by an Euler-Bernoulli beam model. We assume that deflections due to bending are large in comparison to deflections due to shear deformations, which is generally true for long, slender beams. That is to say that we assume that the cross section of the beam is perpendicular to the bending line (i.e. plane sections before bending remain plane after bending).

We define the curvature of a plane curve is the rate of turning of the tangent to the curve, i.e. the angle turned through by the tangent per unit length of curve:

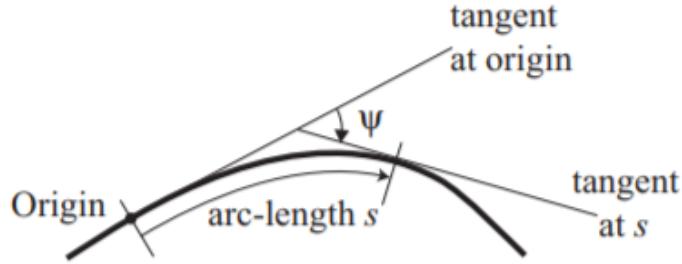


Figure 2: Definition of curvature using a curvilinear coordinate system.

Curvature is

$$\kappa = \frac{d\psi}{ds}, \quad (2)$$

assuming that the deflections and rotations in question are small, then ψ is small, and the curvature can be related to the deflection of a beam v by the following relationship,

$$\begin{aligned} \kappa &= \frac{d\psi}{ds} \approx \frac{d\psi}{dx} \approx \frac{d}{dx} \left(\frac{dv}{dx} \right) \\ &= \frac{d^2v}{dx^2}. \end{aligned} \quad (3)$$

It is known from experiments and can be shown using calculus that the internal bending moment of a beam is proportional to its curvature, assuming linear elasticity.

$$M_{xx} = B\kappa, \quad (4)$$

where the bending stiffness B defined above is an elastic constant of the beam:

$$B = EI_{xx}, \quad (5)$$

where E is the Young's modulus of the material, and I_{xx} the "second moment of area" of the cross-section about it's major axis.

The Young's modulus is a material property that can be easily measured from bend tests of the bulk material of choice. The second moment of area of the beam changes as the cross section of the beam changes. It can be crudely approximated by smearing all the material between the flanges into a central web, so that it looks like an I beam, as shown in Figure 3. Using this approximation, the second moment of area was calculated to be $19,560\text{mm}^4$.

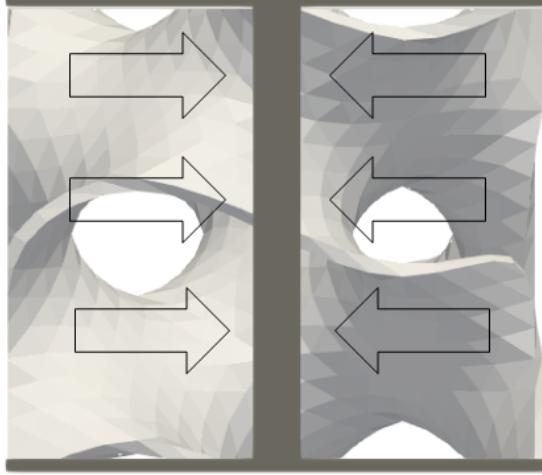


Figure 3: Approximating the second moment of area about the neutral axis by 'smearing' the core material on to a central web.

Then, by integrating the bending moments along the beam, the deflection of the beam as a function of distance from one of the supports $u(x)$ was calculated.

This model of the beam behaviour is likely to under-estimate the deflections of the beam. This model does not take into account shear deflections which are not negligible since the beam is not thin, on the contrary: the depth is fairly large when compared to its length, with a depth/span ratio of 7.1. Furthermore, the bending stiffness of the beam may have been overestimated by smearing all of the material between the flanges onto a central web to calculate the second moment of area: by doing this, the web accounts for a third of the total bending stiffness of the beam. However, this hypothesis has not been tested.

2.2 Finite Element Method

FEM is well documented in literature and the reader is recommended to look at the following booklist for more detail: Cook, R.D. Finite Element Modelling for Stress Analysis, Wiley, (1995).; Ottosen, N.S. & Petersson, H. Introduction to the Finite Element Method, Prentice Hall, 1992.; Cook, R.D. Finite Element Modelling for Stress Analysis, Wiley, (1995).; Fish, J. A First Course in Finite Elements, John Wiley & Sons (2007). A brief summary of the method is written below, and adapted from Section 2 of Girolami et al, The Statistical Finite Element Method (2019) [5].

We assume that the mathematical representation of the deflection of a beam under loading is modelled by $u(x)$. This is governed by partial differential equation $\mathcal{L}_\theta u(x) = -f(x)$ where \mathcal{L}_θ is some linear differential operator with respect to the model parameters θ (e.g. material properties such as bulk material stiffness). Given a test function $v(x)$, it follows from standard variational calculus that the solution $u(x)$ of the partial differential equation has to satisfy the weak form,

$$\mathcal{B}_\theta(u(x), v(x)) = \langle f(x), v(x) \rangle, \quad (6)$$

where \mathcal{B}_θ is the bilinear form corresponding to \mathcal{L}_θ and $\langle \cdot, \cdot \rangle$ is the appropriate Hilbert space inner product. The functions $u(x)$ and $v(x)$ are represented as expansions of basis functions, $u(x) = \sum_{i \in I} u_i \phi_i(x)$ and $v(x) = \sum_{j \in I} v_j \phi_j(x)$, where $\phi_i(x)$ is a basis function. In an FE model, with mesh resolution h , these infinite

expansions are approximated as finite dimensional summations replacing \mathbf{I} with a countable basis index \mathbf{I}_h . For FE models one obtains $\sum_{i,j \in \mathbf{I}_h} u_i \mathcal{B}_\theta(\phi_i(x), \phi_j(x)) = \langle f(x), \phi_j(x) \rangle$ and our FE solution is

$$\mathbf{u}_h = u_h(\mathbf{x}_h) = \mathbf{A}^{-1} \mathbf{b}, \quad (7)$$

where the elements of the matrix above are each $A_{i;j} = \mathcal{B}(\phi_i(x), \phi_j(x))$, $b_j = \langle f(x), \phi_j(x) \rangle$. Given a vector of spacial coordinates \mathbf{x}_u , $\mathbf{u}_h \in \mathbb{R}^{N \times 1}$ is the FE approximation to the function $u(x)$ at spacial coordinates \mathbf{x}_u .

In their paper, 'The Statistical Finite Element Method' [5], Girolami, Cirak, Yin et al. developed an implementation of the Finite Element Method that deals with the particular difficulties of the presence of second derivatives of $u(x)$ and $v(x)$ in the shell bilinear form that requires the approximants to be square integrable. In this implementation, the basis functions are smooth quartic box splines that are defined on triangular meshes with certain symmetry and no boundaries; see Cirak et al. (2000) [8]; Cirak and Ortiz (2001) [7]; Cirak and Long (2011) [7] for further details. It was this implementation that was used to model the deflections. The deflected shape can be seen in Fig. 4.

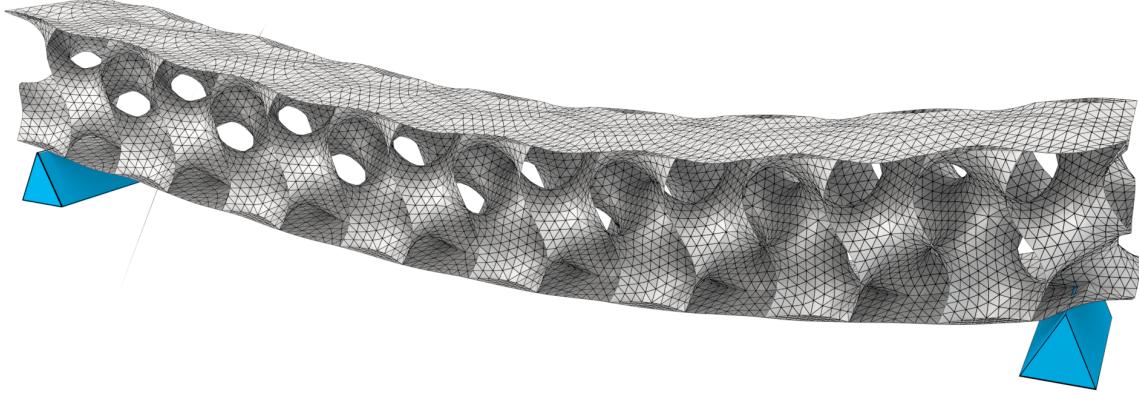


Figure 4: Deflection of the gyroid beam, as approximated by FEM.

3 Manufacturing the gyroid beam

The gyroid beam is a complex geometry that can only be manufactured via additive methods, such as 3D printing. The gyroid beam was 3D printed using the facilities in the Dyson Centre for Engineering Design and tested using the Instron machines in the Fatigue Lab at Cambridge University Engineering Department. The most common choice for 3D printed materials are PLA and ABS. PLA was the material of choice as it is inexpensive, has good mechanical properties and was readily available. Similar size stainless steel, aluminium or titanium beams can be additively manufactured at considerably higher prices. The world's first single-span 3D-printed steel bridge has been printed by The Alan Turing Institute in partnership with the 3D printing company, MX3D [1].

The main challenge with 3D printing the beam was making the shell thin enough so that it could be modelled accurately using thin-shell mechanics. The print settings that were found to work best was a 50% infill setting (although the shell thickness was small enough such that the true infill was actually 100%). In order that the face plates were supported while printing, the beam was printed on its side, as pictured in Fig.5.



Figure 5: the total time to 3D print the gyroid beam was 16 hours, watch the full printing process: <https://www.youtube.com/watch?v=oZDFUwJmiXQ> .

4 Ranking the models using a Bayesian approach

We would like first to assess which of the two models, FEM or Euler-Bernoulli, best explains the data. In order to quantitatively rank the success of these models in predicting deflection behaviour of the beam, we will employ a stochastic model of the data using Bayesian statistics to investigate the likelihood of our deflection data given a particular model.

We assume that the mathematical representation of the process, the deflection of the beam under loading, is modelled by $u(x; \boldsymbol{\theta})$. The Euler-Bernoulli model has parameters,

$$\boldsymbol{\theta} = \{W, E, I_{xx}\}, \quad (8)$$

where W is the assumed loading applied to the model, E is the assumed Young's Modulus and I_{xx} is the assumed second moment of area of the beam about its principal axis. For the FEM model, these parameters will be different, but also assumed to be known.

We will look at the Gaussian distribution from a Bayesian point of view. It is first assumed that the deflection data is normally distributed about a known mean,

$$y_i \sim \mathcal{N}(\mu, \sigma^2), \quad (9)$$

where

$$\mu = u(x_i; \boldsymbol{\theta}) = u_i. \quad (10)$$

Given that u is fixed, then the conjugate prior for σ^2 is an inverse gamma distribution,

$$\sigma^2 \sim IG(\alpha, \beta). \quad (11)$$

Doing a search over the parameters α and β for the maximum marginal likelihood will give us an idea of how well the data fits the model, and so how successful the model is. It may also tell us something about σ^2 .

The mathematical derivation for the marginal likelihood of our data,

$$L = P(\mathbf{y}|\mathbf{u}, \alpha, \beta) = \prod_i P(y_i|u_i, \alpha, \beta), \quad (12)$$

is written in the appendix, Section 8.1.

4.1 Experiment 1: deflection data over different loads

To obtain the data, the deflection of the beam across the span, a laser system and reflective tape were used. The laser system measured the distance between two reflective tapes, one attached to the beam, and one attached to the Instron machine, as pictured. The deflected distances were subtracted from initial distances to get a deflection measurement.

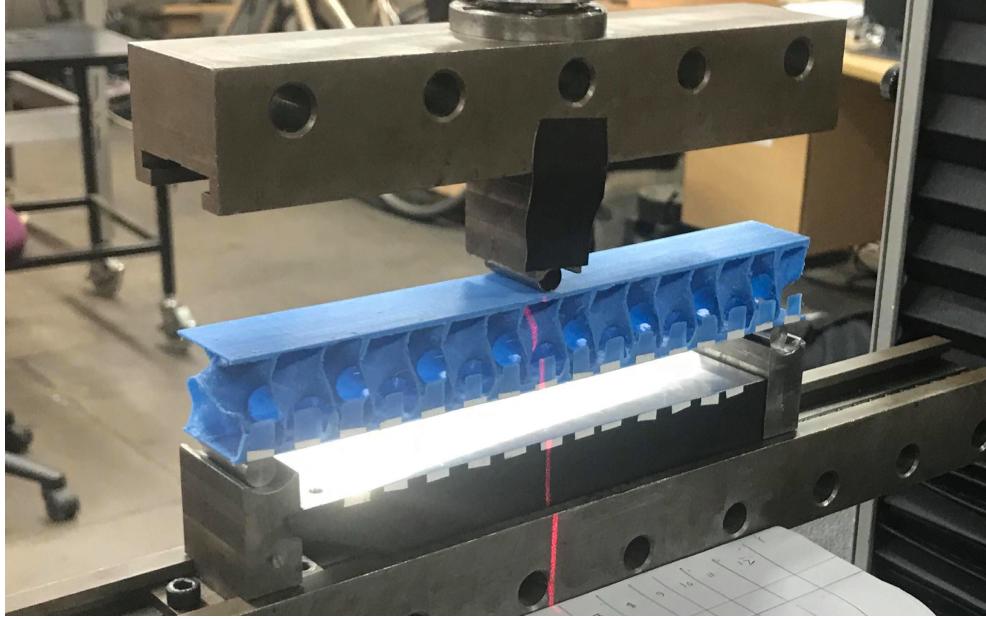


Figure 6: Performing a 3-point bend test on the gyroid beam using an Instron machine. The laser measures the distance between the reflective tape placed on the bottom of the beam and the Instron machine. Some uncertainty in the measurement was introduced by rotating the laser round to take readings across the beam.

The primary sources of observation error were that the tapes were not perfectly planar or their edges perfectly parallel and laser was moved around on a tripod to point it at each set of tapes. Only one set of readings were taken for this experiment.

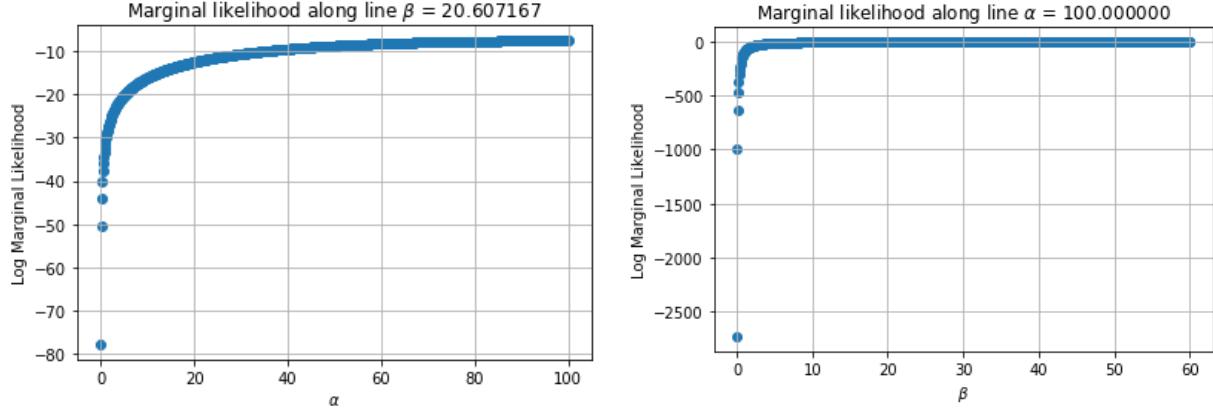
The deflections were measured at loads of 200N, 250N and 300N. A plot of the deflections at the different loads is shown in Fig.?? -??. The 200N deflection measurements look to be noisy when compared to the 250N and 300N deflection. This is because the relative magnitude of the measurement error versus deflection is higher at lower deflections, and the measurement error had lowered for the 250N as we had taken more care with the measurements as we got used to the apparatus. It should be noted that the 300N load caused a small amount of plastic deformation, as deflection was outside of the linear-elastic regime.

4.2 Use of Marginal Likelihood to rank the Euler-Bernoulli Model against the Finite Element Model

We calculate the marginal likelihood of the data to optimise α and β . This is known as the maximum likelihood (ML) estimate of α and β . Assuming that α and β can take any value in the ranges,

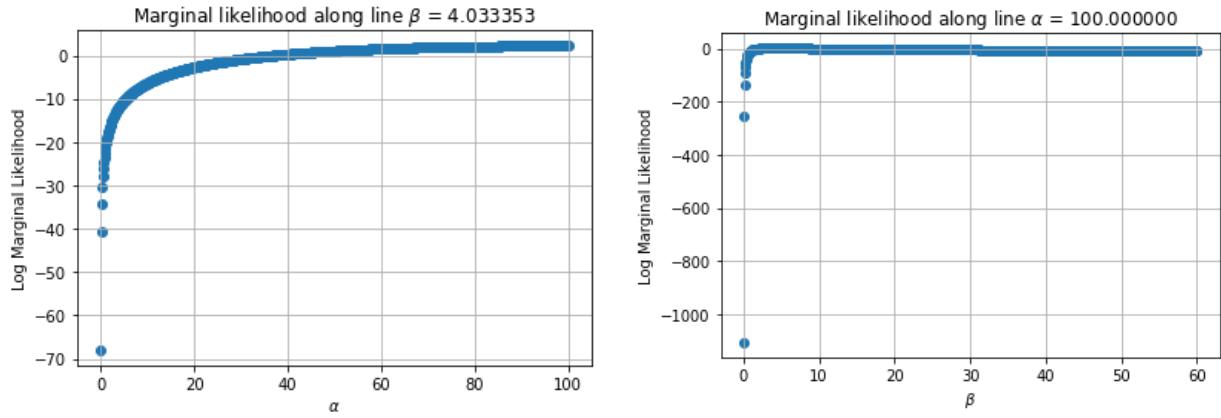
$$\begin{aligned} 0 \leq \alpha &\leq 100 \\ 0 \leq \beta &\leq 60 \end{aligned}$$

Fig. 7 shows the marginal likelihood plots for the Euler Bernoulli model and Fig.?? for the FEM model.



(a) Log marginal likelihood as a function of α , with β optimum.
(b) Log marginal likelihood as a function of β , with α optimum.

Figure 7: Log marginal likelihood plots showing the log likelihood of the data given the Euler-Bernoulli model and parameters α and β .



(a) Log marginal likelihood as a function of α , with β optimum.
(b) Log marginal likelihood as a function of β , with α optimum.

Figure 8: Log marginal likelihood plots showing the log likelihood of the data given the FEM model and parameters α and β .

Looking qualitatively at the deflection results of Fig.?? - ??, it can be seen that the FEM model tends to over-estimate the deflection, although it does pick up the general shape of the deflection fairly well. The Euler-Bernoulli beam model, on the other hand, tends to under-estimate the deflection, as expected (see Section 2.1).

Given this mismatch between the model and the data, it means that the error, $\varepsilon_i = y_i - u_i$, is not IID (independent, identically distributed), and so the basic statistical model from Section 4 cannot explain the data. In turn this means that the statistical model has low identifiability and the marginal likelihood will not converge on a value of α . This is shown to be the case in Fig.7 for the Euler Bernoulli model and in Fig.?? for the FEM model.

The maximum likelihood was found at the edge of the search at $\alpha = 100$, and continues to rise for values $\alpha > 100$, in other words, there is no maximum likelihood estimate of α . It would be incorrect to select an arbitrary value of α and β to compare the likelihood of each model, so we shall put a uniform prior distribution on the parameters α and β ,

$$\pi(\alpha, \beta | \kappa) = \kappa = \frac{1}{(\alpha_{max} - \alpha_{min})(\beta_{max} - \beta_{min})} \quad \alpha_{min} < \alpha < \alpha_{max}, \beta_{min} < \beta < \beta_{max}, \quad (13)$$

where κ is a hyper-parameter (a known constant in this case), to marginalise out α and β to obtain the marginal likelihood,

$$P(\mathbf{y}|\mathbf{u}) = \int \int P(\mathbf{y}|\mathbf{u}, \alpha, \beta) \pi(\alpha, \beta | \kappa) d\alpha d\beta. \quad (14)$$

Since we have sampled α and β over a discrete grid, then this integral can be approximated by

$$P(\mathbf{y}|\mathbf{u}) \approx \frac{1}{N} \sum_{i,j} P(\mathbf{y}|\mathbf{u}, \alpha_i, \beta_j), \quad (15)$$

where N is the number of samples in the grid. The results of this value for the 2 models and 3 sets of data is shown in Table 1.

Table 1: $P(\mathbf{y}|\mathbf{u})$ values for each model and data set

\mathbf{y}	FEM	Euler-Bernoulli Model
$\mathbf{y}_{W=200N}$	0.1718	0.0022
$\mathbf{y}_{W=250N}$	0.3749	0.0001
$\mathbf{y}_{W=300N}$	0.0912	0.0000

If we assume a uniform prior for each model $p(\mathbf{u}_i)$, then the posterior for each model follows as,

$$P(\mathbf{u}_i | \mathbf{y}) = \frac{P(\mathbf{y}|\mathbf{u}_i)}{\sum_j P(\mathbf{y}|\mathbf{u}_j)}. \quad (16)$$

The ranking for each model is shown in Table 2 for the 2 models and 3 sets of data.

Table 2: $P(\mathbf{u}|\mathbf{y})$ values for each model and data set

\mathbf{y}	FEM	Euler-Bernoulli Model
$\mathbf{y}_{W=200N}$	0.9874	0.0126
$\mathbf{y}_{W=250N}$	0.9997	0.0003
$\mathbf{y}_{W=300N}$	0.9998	0.0002

Table 2 shows that this statistical model of the data has overwhelmingly ranked the FEM model with higher likelihood, and therefore the better model, for all 3 data sets. As the loading, W , increased, the error between the data and the Euler-Bernoulli model increased more than the error between the data and the FEM model. This is reflected in Table 2, as the posterior of the FEM model increases as the loading is increased, and the posterior for the Euler-Bernoulli model decreases as the loading is increased. All further analysis will be done on the FEM model alone.

5 A more sophisticated statistical model

The previous analysis did not take into account a mismatch error between the model and reality, resorting to using a Gaussian noise to completely model the error between the observed values, y and the model u . The following analysis extends the stochastic model of the data to use an additive functional error model to represent the model-reality mismatch, as presented in the paper 'The Statistical Finite Element Method' [5].

5.1 Probabilistic representation of FEM

In their paper, 'The Statistical Finite Element Method' [5], Girolami et al. induce uncertainty in the Finite Element model due to incomplete knowledge of the right hand side forcing term $f(x)$ (see Section 2.2), and uncertain parameters, $\boldsymbol{\theta}$. These uncertainties are formally incorporated into the probabilistic representation of the FE method by defining a probabilistic model of the randomised forcing term and making the linear differential operator, $\mathcal{L}_{\boldsymbol{\theta}}$, a randomised operator with respect to the uncertain parameters, $\boldsymbol{\theta}$. However, for simplicity we employ the traditional representation of the finite element model, given certain parameters, $\boldsymbol{\theta}$, and right hand side forcing term, $f(x)$.

5.2 Finite Element Model Conditioned on Data

Consider a vector of measured deflections denoted as $\mathbf{y} \in \mathbb{R}^{N \times 1}$. This N dimensional vector represents N measurements taken simultaneously from N sensors. The important point to note here is that a statistical model of the data posits a true underlying generating process.

In their seminal work Kennedy and O'Hagan (2001) employed an additive functional error model to represent the model-reality mismatch, such that the 'true' partially known generating process values at the observation points are $\boldsymbol{\eta} \in \mathbb{R}^{N \times 1}$. As $\boldsymbol{\eta}$ is unobserved, its value is considered as a random function. This forms an additive regression model with random mismatch error, $\boldsymbol{\delta}_h \in \mathbb{R}^{N \times 1}$, such that $\boldsymbol{\eta} = \mathbf{u}_h + \boldsymbol{\delta}$, with $\boldsymbol{\delta} \sim \mathcal{N}(0, \mathbf{K}_{\lambda})$. Here, $\mathbf{K}_{\lambda} \in \mathbb{R}^{N \times N}$ is a matrix of covariance function evaluations $k_{\delta}(x, x')$, with length-scale λ , evaluated at the points \mathbf{x}_y ,

$$K_{ij} = \exp(-\lambda|x_i - x_j|^2). \quad (17)$$

The final component in the data model is the measurement error, which in this presentation is also assumed Gaussian. In which case, the data is now defined as $\mathbf{y} = \boldsymbol{\eta} + \boldsymbol{\epsilon}$, with $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma_{\epsilon}^2 \mathbf{I})$. This is to say that the measurement error is assumed to be IID (independent, identically distributed), i.e. not dependent of the position of the measurement across the beam. The conditional probability of the data under random vectors of the measure of the FE model-data mismatch, gives $p(\mathbf{y}|\mathbf{u}_h, \boldsymbol{\theta}, \boldsymbol{\delta}, \lambda, \sigma_{\epsilon}) = \mathcal{N}(\mathbf{u}_h + \boldsymbol{\delta}, \mathbf{I}\sigma_{\epsilon}^2)$. The marginal probability of the data, $p(\mathbf{y}|\boldsymbol{\theta}, \boldsymbol{\delta}, \lambda, \sigma_{\epsilon})$, where the random vector $\boldsymbol{\delta}$ can be integrated out, and known parameters $\boldsymbol{\theta}$ are omitted in the notation for brevity, follows as,

$$p(\mathbf{y}|\lambda, \sigma_{\epsilon}) = \mathcal{N}(\mathbf{u}_h, \mathbf{K}_{\lambda} + \sigma_{\epsilon}^2 \mathbf{I}). \quad (18)$$

5.3 Using Monte Carlo procedures

We would like to infer information about the 'true' values of the parameters λ and σ_{ϵ} . We could calculate the posterior distribution, $p(\lambda, \sigma_{\epsilon}|\mathbf{y})$, analytically, but we will be using Monte Carlo procedures to generate the posterior distribution. In particular we will be using the Metropolis algorithm.

5.3.1 Metropolis Algorithm

The Metropolis Algorithm is as follows, using vector notation for brevity, $\mathbf{w}_s = [\sigma_{\epsilon}, \lambda]^T$:

Algorithm 1 Metropolis Hastings Algorithm

```

1: procedure GETSAMPLES( $S$ )                                 $\triangleright$  Metropolis algorithm,  $S$  is number of iterations
2:    $s = 0$ 
3:   Choose  $\mathbf{w}_s$ 
4:   while  $s \leq S$  do
5:      $s = s + 1$ 
6:     Generate  $\tilde{\mathbf{w}}_s$  from  $p(\tilde{\mathbf{w}}_s | \mathbf{w}_{s-1})$ 
7:     Compute acceptance ratio  $r$ 
8:     Generate  $u$  from  $U(0, 1)$                                  $\triangleright$  Uniform distribution
9:     if  $u \leq r$  then
10:      Accept sample,  $\tilde{\mathbf{w}}$ 
11:       $\mathbf{w}_s = \tilde{\mathbf{w}}_s$ 
12:    else
13:       $\mathbf{w}_s = \mathbf{w}_{s-1}$ 
14:   return accepted samples

```

Generating $\tilde{\mathbf{w}}_s$: $\tilde{\mathbf{w}}_s$ is generated from a Gaussian proposal distribution, located at \mathbf{w}_{s-1} , which is the last accepted sample. $p(\tilde{\mathbf{w}}_s | \mathbf{w}_{s-1}) = \mathcal{N}(\mathbf{w}_{s-1}, \Sigma)$. Σ is the covariance of the Gaussian proposal density. The Gaussian is a popular choice for the proposal density because it is symmetric: moving to $\tilde{\mathbf{w}}_s$ from \mathbf{w}_{s-1} is just as likely as moving from \mathbf{w}_{s-1} to $\tilde{\mathbf{w}}_s$, which simplifies the acceptance ratio,

$$r = \frac{p(\tilde{\mathbf{w}}_s | \mathbf{y})}{p(\mathbf{w}_{s-1} | \mathbf{y})} \frac{p(\mathbf{w}_{s-1} | \tilde{\mathbf{w}}_s, \Sigma)}{p(\tilde{\mathbf{w}}_s | \mathbf{w}_{s-1}, \Sigma)} = \frac{p(\tilde{\mathbf{w}}_s | \mathbf{y})}{p(\mathbf{w}_{s-1} | \mathbf{y})}. \quad (19)$$

Using a nonsymmetric proposal density is known as the Metropolis-Hastings algorithm, whereas using a symmetric proposal density is known as the Metropolis algorithm. What's left is the first term, the ratio of the posterior density at the proposed sample to that of the old sample. We do not know the posterior distributions exactly because we cannot normalise them. However, because we are interested in a ratio, the normalisation constants cancel. So, applying Bayes rule, we can substitute the ratio of posteriors with the ratio of priors multiplied by the ratio of likelihoods. Putting independent priors on σ_ϵ and λ , $\pi(\sigma_\epsilon)$ and $\pi(\lambda)$, leads us to the expression,

$$r = \frac{p(\mathbf{y} | \mathbf{u}_h, \sigma_{\epsilon,s}, \lambda_s)}{p(\mathbf{y} | \mathbf{u}_h, \sigma_{\epsilon,s-1}, \lambda_{s-1})} \frac{\pi(\sigma_{\epsilon,s})\pi(\lambda_s)}{\pi(\sigma_{\epsilon,(s-1)})\pi(\lambda_{s-1})}. \quad (20)$$

Since σ_ϵ and λ are strictly positive, the exact priors that we are putting on them are Gamma distributions,

$$\pi(\sigma_\epsilon | \alpha_{\sigma_\epsilon}, \beta_{\sigma_\epsilon}) = Ga(\alpha_{\sigma_\epsilon}, \beta_{\sigma_\epsilon}), \quad (21)$$

$$\pi(\lambda | \alpha_\lambda, \beta_\lambda) = Ga(\alpha_\lambda, \beta_\lambda), \quad (22)$$

where the gamma distribution pdf for some random variable, z is defined as,

$$P(z | \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} z^{\alpha-1} \exp(-\beta z). \quad (23)$$

The prior allows us to express a belief about the true values of σ_ϵ , λ . σ_ϵ is likely to be small but non zero, since we would expect a small percentage measurement error, and being a length scale, λ will be very small, e.g. $\lambda \approx 10^{-3}$, if δ is going to be a non-zero vector. We might then pick a gamma distribution with a high likelihood close to 0. This can be done by setting the shape parameter $\alpha_{\sigma_\epsilon} = 1.0$, $\alpha_\lambda = 1.0$. The rate parameters, β_{σ_ϵ} and β_λ should depend on the relative sizes of σ_ϵ and λ . A high rate parameter will give a smaller spread. A low value of λ is necessary for δ to have a visible effect on the data. Since λ is going to be much smaller than σ_ϵ , then it should have a lower spread about 0. Values of $\beta_{\sigma_\epsilon} = 1.0$, $\beta_\lambda = 100$ were used. The likelihood is calculated from equation 18. An illustrative example of using this method on example data is shown in Section 8.2.1 of the appendix. At this point we also perform a change of variables on λ ,

$$\lambda = \exp(\zeta), \quad (24)$$

since λ is very small and when sampling from the posterior using the Metropolis algorithm, the sample rate is small. For the full derivation of the resulting prior on ζ , see Section 8.2.2.

5.3.2 Multiple sets of observations

Using multiple sets of sensor data will reduce the dependence of the posterior on the prior, and will result in a target distribution that has higher likelihood around the 'true' values of ζ and σ_ϵ .

In the case where there are multiple sets of observations \mathbf{y}_i at the coordinates x_y , where the sensors are located, for $i = 1, \dots, L$, it can be shown that the joint distribution of all \mathbf{y}_i conditional on \mathbf{u}_h is given in product form by assuming independence between the sets of observations. Let $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_L] \in \mathbb{R}^{N \times L}$, the joint distribution of \mathbf{y}_i conditional on \mathbf{u}_h given all sets of observations is

$$p(\mathbf{Y}|\mathbf{u}_h, \sigma_\epsilon, \zeta) = \prod_i^L p(\mathbf{y}_i|\mathbf{u}_h, \sigma_\epsilon, \zeta), \quad (25)$$

so that the acceptance ratio becomes

$$r = \frac{p(\mathbf{Y}|\mathbf{u}_h, \sigma_{\epsilon,s}, \zeta_s)}{p(\mathbf{Y}|\mathbf{u}_h, \sigma_{\epsilon,s-1}, \zeta_{s-1})} \frac{\pi(\sigma_{\epsilon,s})\pi(\zeta_s)}{\pi(\sigma_{\epsilon,(s-1)})\pi(\zeta_{s-1})}. \quad (26)$$

The result of adding more sets of sensor data is to reduce the size of the posterior distribution to concentrate more around the 'true' values of σ_ϵ and ζ . A MAP (maximum a posteriori) estimate can then be chosen from the posterior distribution. An illustrative example of using multiple sets of observations, using example data, is shown in Section 8.2.3 of the appendix.

5.4 Experiment 2: multiple sets of readings of deflection data for a single load

Following experiment 1 (see Section 4.1), further measurements were taken to get repeat readings of the deflection at 250N. In total 8 sets of readings were taken, the average, $\bar{y} = \sum_{i=1}^8 y_i$, is shown in Fig. 9, the differences between average sensor reading and FEM model, $\bar{y} - u_{FEM}$, are shown in Fig. 10 and the variance of the readings, $VAR(y)$, are shown in Fig. 11.

When compared to the first results, shown in Fig. ??, the deflections are larger. This may be because of a possible crack defect in the beam which could explain the sudden jump in deflection at approximately 40N, as shown in Fig. 12

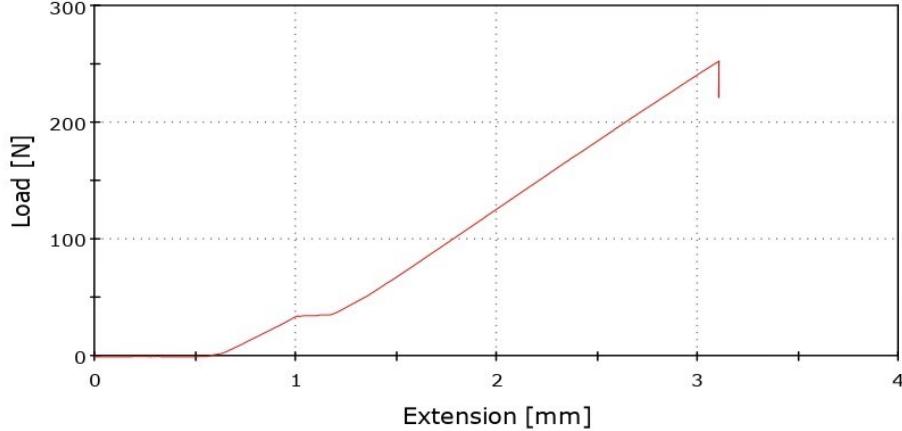


Figure 12: Experiment 2 load-extension curve for the Instron test at W=250N. Note the initial displacement at 0N is due to machine compliance and was not measured by the sensors. Also note the drop off in load once loading was stopped at 250N, this is due to stress relaxation, however, the deflection remains unchanged.

Fig. 12 shows the readings from the 12th position along the beam (leftmost is 1st position) to be erroneous, with much larger deflection readings than expected. This may have been because the laser was picking up reflections from a foreign object, or the tape had been moved from its original measured position during the experiment. For this reason, the readings from the 12th position along the beam has been identified as anomalous and have been omitted from further analysis.

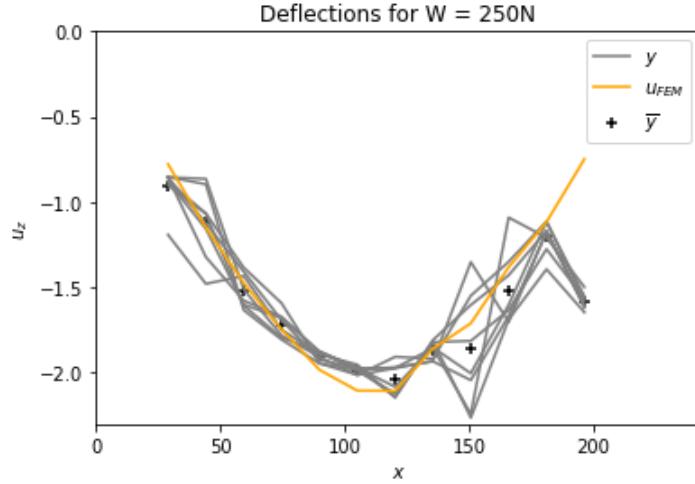


Figure 9: Deflection plot for the gyroid beam at $W=250N$. The orange line is the FEM model, the grey lines are the measured deflections from the Instron test, y , and the black crosses are the mean values, \bar{y}_{bar} .

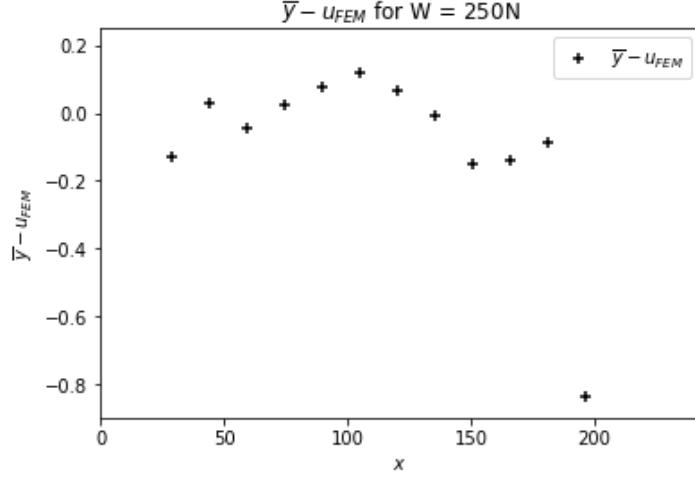


Figure 10: Differences between average sensor reading and FEM model across the beam.

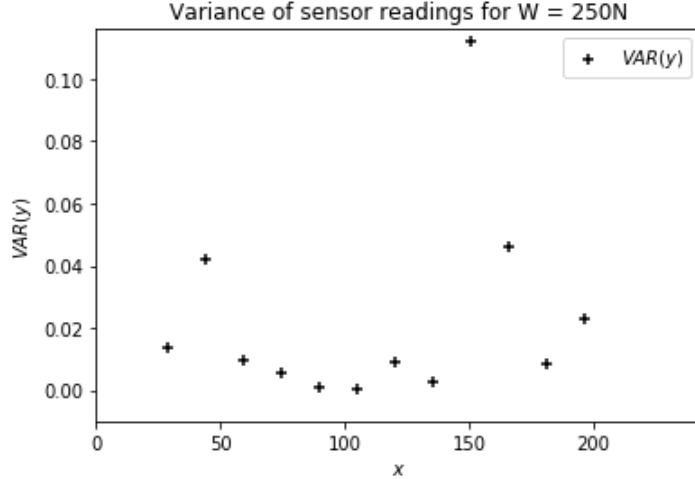


Figure 11: Variance in the sensor readings across the beam.

The variance in y readings on the 2nd, 9th and 10th positions along the beam are high when compared to the rest of the measurements. If the IID assumption was correct (see section 5.2), then in theory, we could pool the data from the different samples (that is, different distances along the beam) as they are all taken from the same population. Pooling the samples in this way, then the sample variance of all the samples with the anomalous readings from the 12th position omitted is $VAR(y) = 0.020181122$. The sample variance with the readings from the 12th, 2nd, 9th and 10th positions omitted is $VAR(y) = 0.005594012$. This shows that the measurement error is not IID. These values have been omitted from parts of the analysis at the cost of having significantly fewer measurements along the beam.

We would expect the value of the MAP estimate of σ_ϵ (with the readings from the 12th, 2nd, 9th and 10th positions omitted) to be close to the empirical, pooled sample standard deviation of $\sigma_{sample} = \sqrt{VAR(y)} = 0.0748$ (to 3 s.f.).

5.5 Use of MCMC to sample the parameters

The MCMC sampler described in Section 5.3.1 was used on the observed data from Experiment 2 to sample from the posterior distribution of the parameters. Fig. 13 shows the samples obtained, with anomalous data from position 12 omitted. Fig. 14 shows the samples obtained with further data from positions 2, 9 and 10 omitted, due to them not being IID. Fig. 14 shows the values of ζ and σ_ϵ to converge, whereas Fig. 13 does not show convergence.

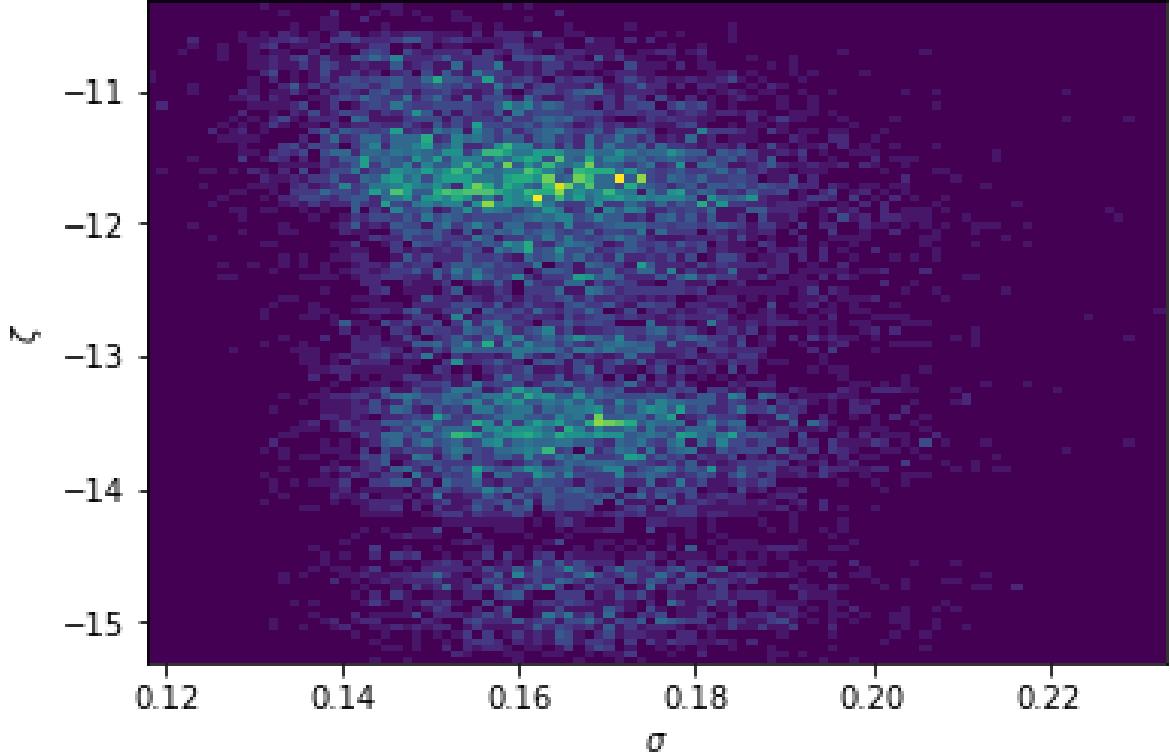


Figure 13: Histogram of 18,160 samples with 45.4% acceptance rate. Using these samples, convergence on neither ζ nor σ_ϵ was found.

This suggests the measurement error, ϵ , is explained by $\sigma_\epsilon = 0.0827$ and that the mismatch error, δ , can be explained by $\zeta = -10.85$. Given that the empirical measurement of the standard deviation of the data, $\sigma_\epsilon = 0.0748$, is reasonably likely value given the posterior distribution, the posterior distribution does not look unreasonable.

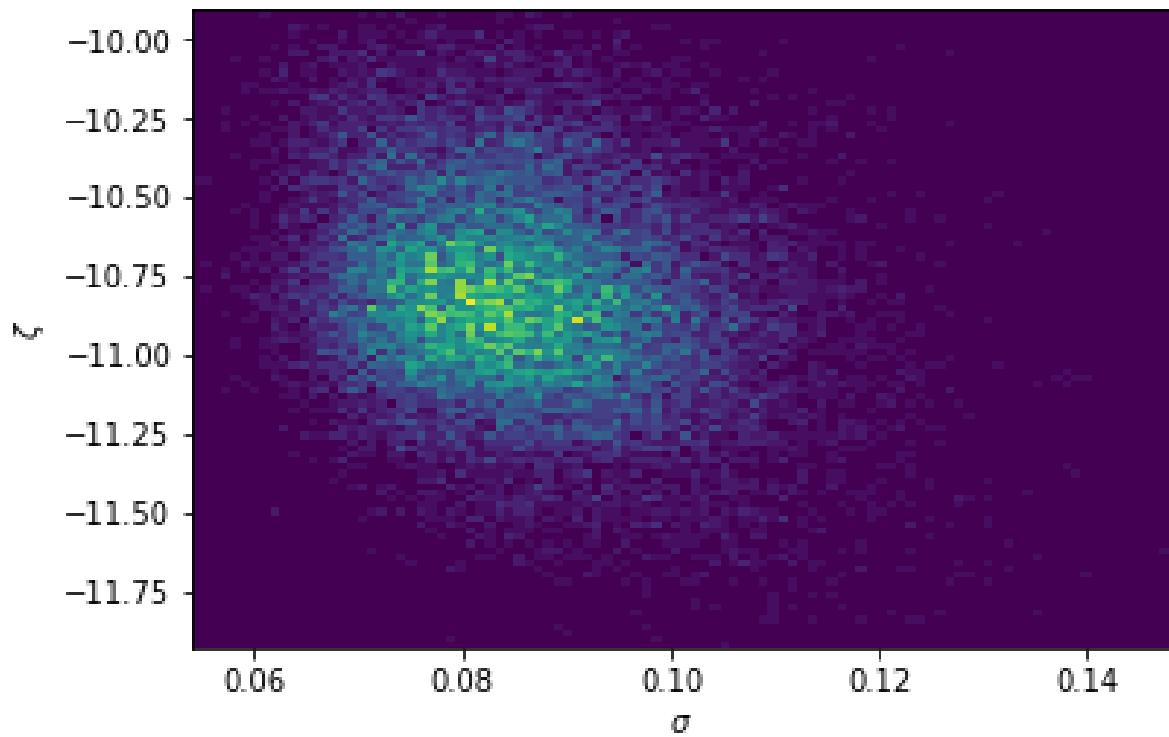


Figure 14: Histogram of 21,093 samples 35.2% acceptance rate. Using these samples, convergence on ζ and σ_ϵ was found, with a small posterior distribution that is roughly aligned with the axes.

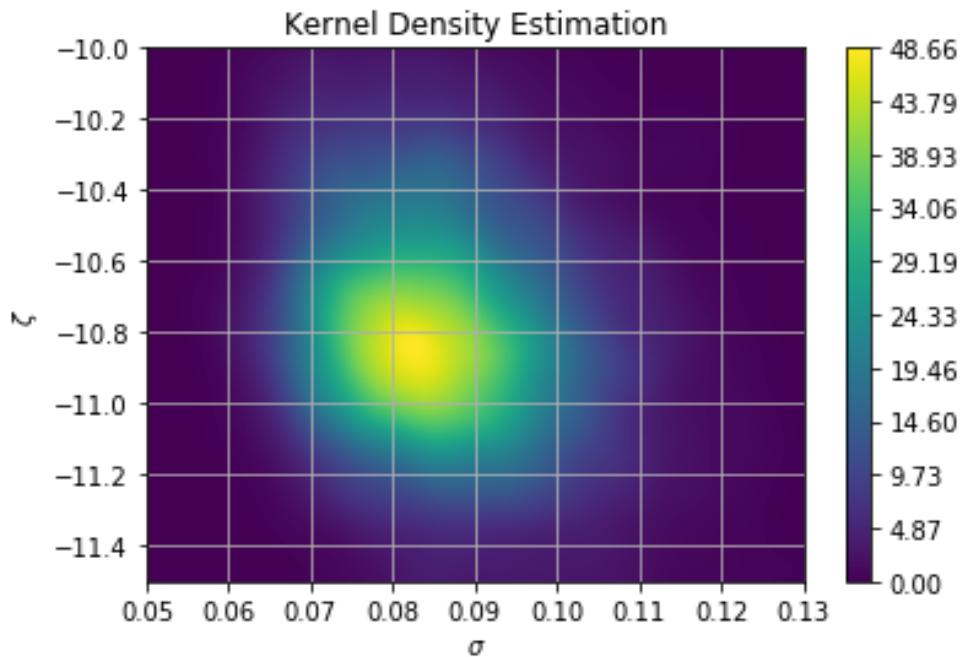


Figure 15: A kernel density estimate of the histogram gives the MAP estimates of $\zeta = -10.85$ and $\sigma_\epsilon = 0.0827$.

6 Discussion

In Experiment 1, the deflections were measured at loads of 200N, 250N and 300N. The 200N deflection measurements, as shown in Fig. ??, look to be noisy when compared to the 250N measurements, as shown in Fig. ??, and 300N measurements, Fig. ???. This is because the relative magnitude of the measurement error versus deflection is higher at lower deflections, and the measurement error had lowered for the 250N as we had taken more care with the measurements as we got used to the apparatus.

The Euler-Bernoulli model underestimated the deflections. The reason identified for this was that the Euler-Bernoulli model was too stiff, and it did not account for shear deflections. The FEM model overestimated the deflections, but was identified in Section 4 as the model with the largest posterior probability, $P(\mathbf{u}|\mathbf{y})$, given all 3 different data sets at 3 different loads. Furthermore, as the load increased, the posterior, $P(\mathbf{u}|\mathbf{y})$ of the FEM model increased, whilst the posterior for the Euler-Bernoulli model decreased. This was reflected in the data, since as the load increased the data moved further and further away from the Euler-Bernoulli model. For this reason, the FEM model was chosen as the better model.

However, the simple stochastic model in Section 4 did not well explain the data. It assumed that the data was normally distributed around the model deflections with some variance. Experiment 2 was designed to identify the model-reality mismatch that the analysis in Section 4 could not identify.

Section 5 employed an additive functional error model to represent the model-reality mismatch, and used Monte Carlo procedures to identify the parameters that described the mismatch under the assumption that the measurement noise was IID (independent, identically distributed). Some of the measurements (positions 2, 9, 10, where the leftmost is position 1) showed that this assumption was incorrect, as they had a higher variance than the other measurements. This meant that the posterior distribution did not pick out correct MAP values of sigma, as shown in Fig. 13.

With the results that did not follow the IID assumption omitted, Fig. 14 shows the stochastic model was able to pick out MAP estimates of these parameters of $\sigma_\epsilon = 0.0827$ and $\zeta = -10.85$. The inferred MAP estimate $\sigma_\epsilon = 0.0827$ is close to the empirical estimate of the standard deviation of the pooled data, $\sigma_{sample} = 0.0748$. With the σ_{sample} also being a likely value in the generated posterior distribution, this indicates success in identifying the correct posterior density.

To qualitatively assess how good our model is at describing the mismatch error, δ , we could compare the measured data from experiment 2, as shown in Fig. 16, to simulated data, generated with the inferred parameters, $\sigma_\epsilon = 0.0827$ and $\zeta = -10.85$, as shown in Fig. 17.

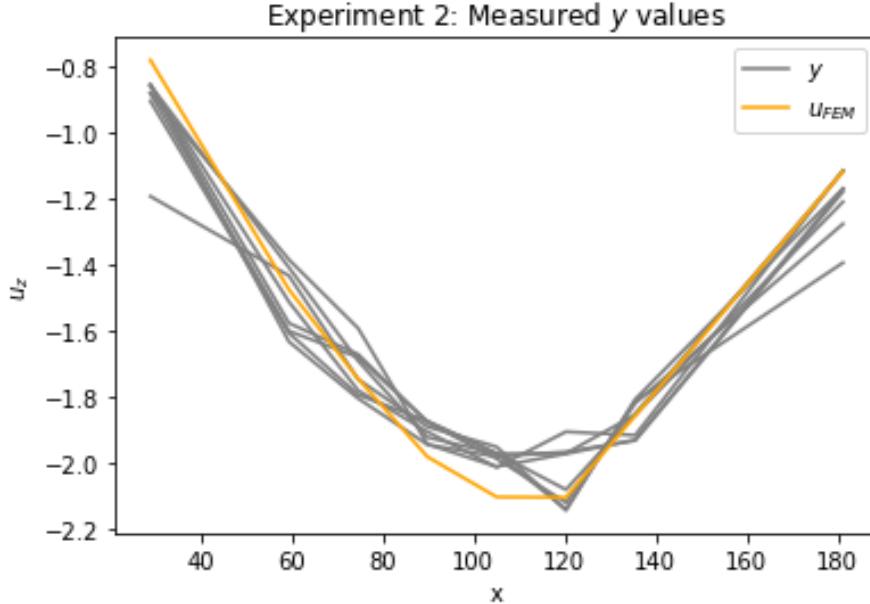
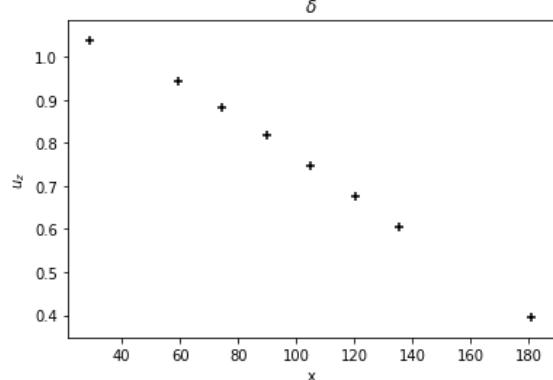
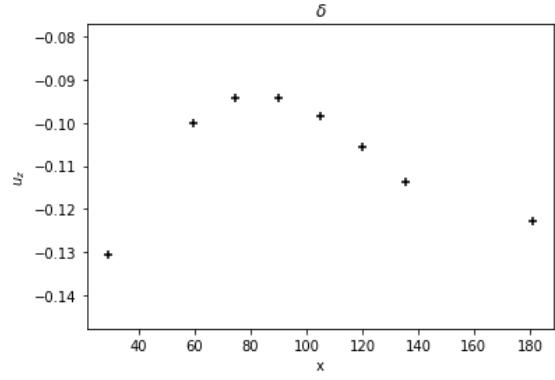
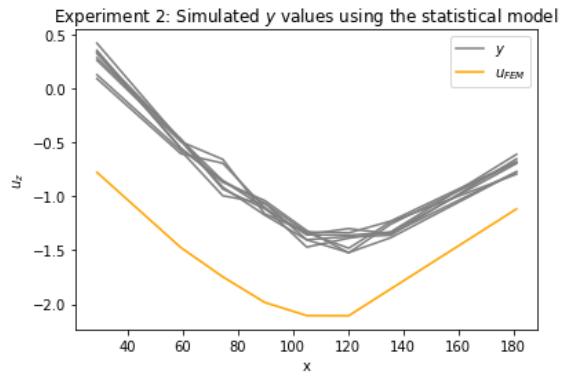
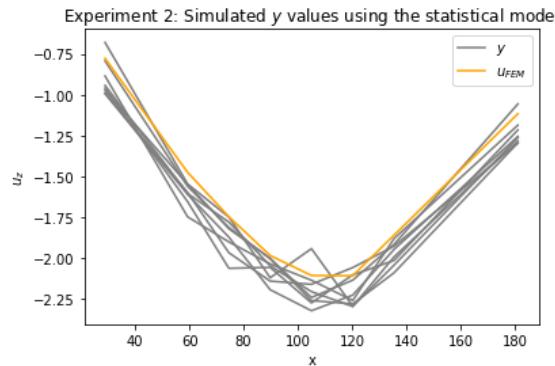
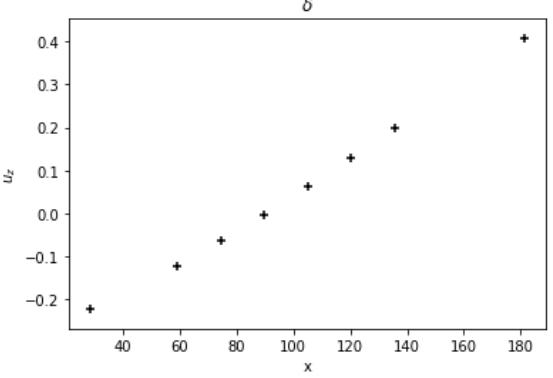
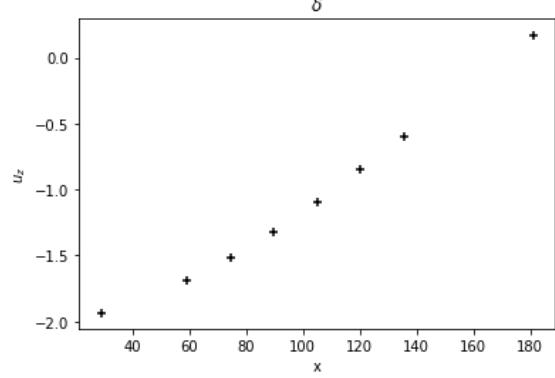
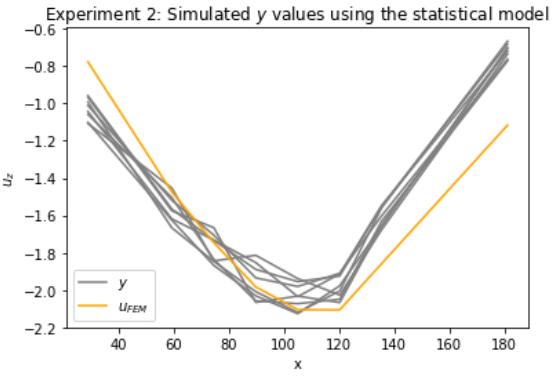
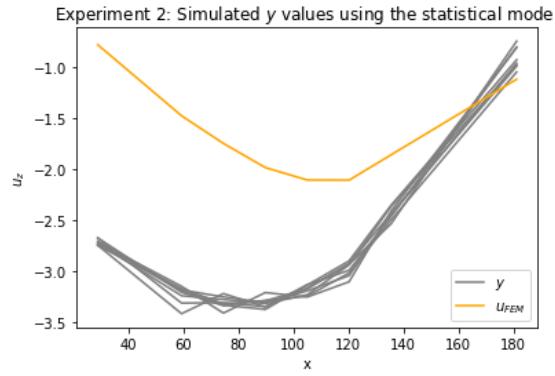


Figure 16: Measured deflection readings from experiment 2, y , with values from positions 2, 9, 10 and 12 omitted, and the FEM model.



(a) Sample no. 1.

(b) Sample no. 2.



(c) Sample no. 3.

(d) Sample no. 4.

Figure 17: Simulated data samples generated with parameter values, $\sigma_\epsilon = 0.0827$ and $\zeta = -10.85$. For each sample, 1 to 4, on the top figure shows the generated data, y , and the bottom figure shows the generated mismatch error, δ .

Comparing at Fig. 17 to Fig. 16, it can be seen that the magnitude of the mismatch error is much too large. We can fix this by adding another parameter into our model, a scale factor, σ_δ , on the mismatch error, δ , such that the 'true' deflection values are $\eta = \mathbf{u}_h + \delta$, with,

$$\delta \sim \mathcal{N}(0, \sigma_\delta^2 \mathbf{K}_\lambda) \quad (27)$$

Monte Carlo procedures, such as the one described in Section 5.3, can be used to find a MAP estimate of the parameter σ_δ . We will first fix σ_ϵ to its MAP estimate of $\sigma_\epsilon = 0.0827$. Putting a gamma prior on σ_δ , we will do a 2D search over parameters σ_δ and ζ .

The MCMC sampler described in Section 5.3.1 was used on the observed data from Experiment 2 to sample from the posterior distribution of the parameters. Fig. 14 shows the values of ζ and σ_δ to converge.

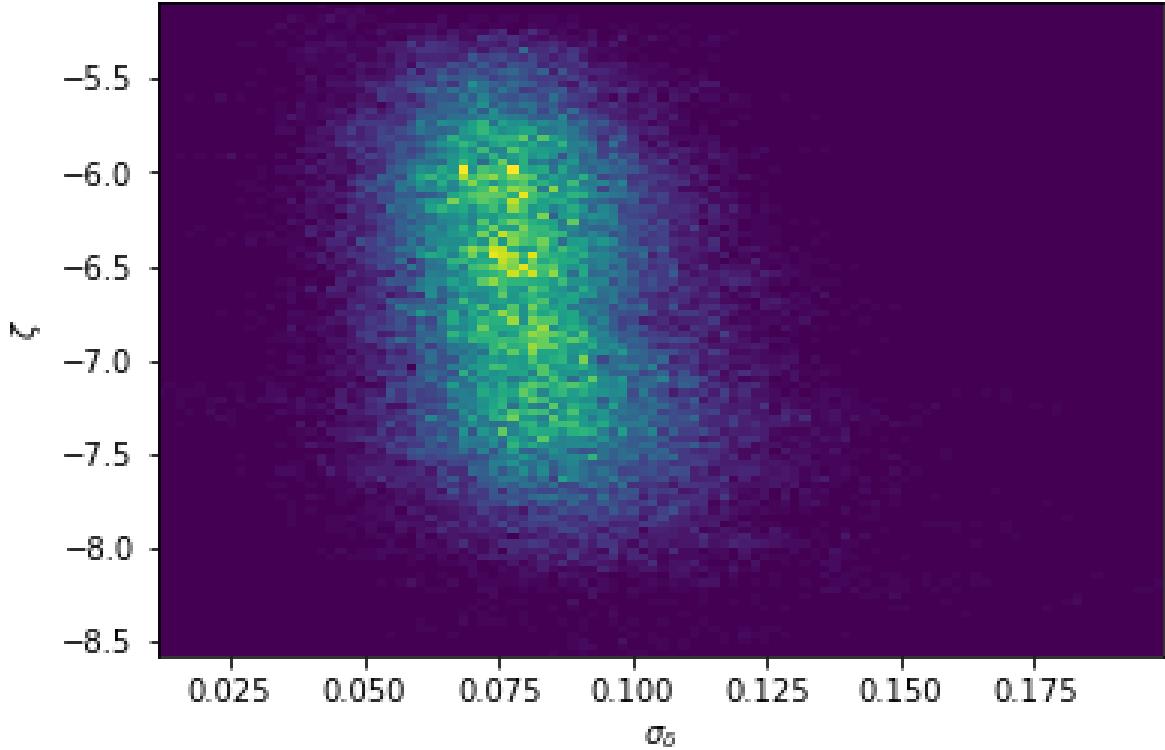
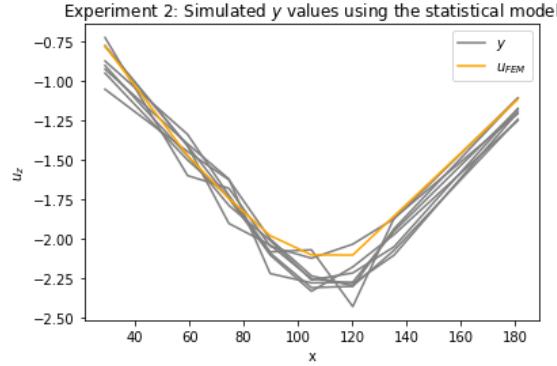
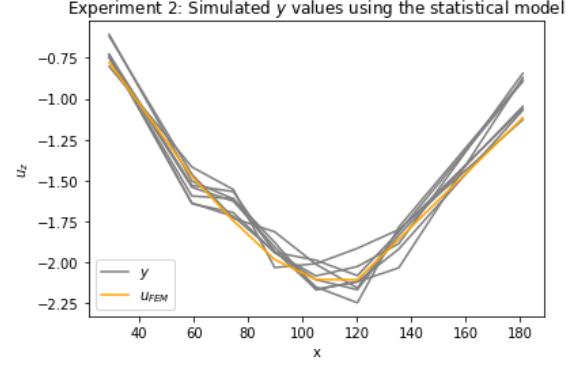


Figure 18: Histogram of 56,963 samples with 81.4% acceptance rate. The kernel density estimate gives MAP estimates of $\sigma_\delta = 0.077$ and $\zeta = -6.45$.

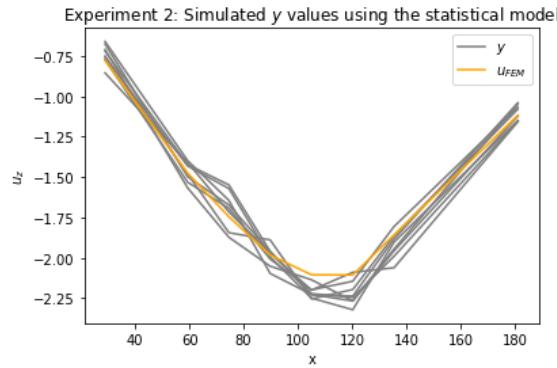
Again, qualitatively assessing how good our model is at describing the mismatch error, δ , we will simulate data using MAP estimates of $\sigma_\delta = 0.077$ and $\zeta = -6.45$, as shown in Fig. 19.



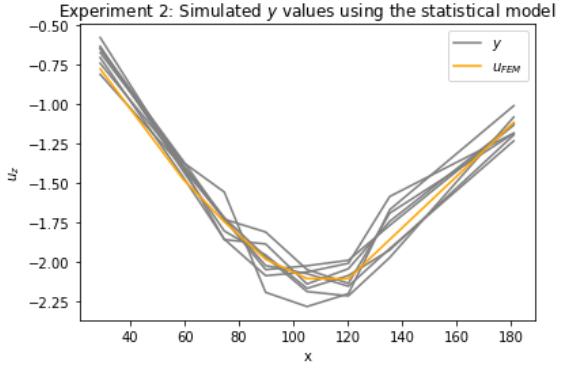
(a) Sample no. 1.



(b) Sample no. 2.



(c) Sample no. 3.



(d) Sample no. 4.

Figure 19: Simulated data samples generated with parameter values, $\sigma_\epsilon = 0.0827$, $\sigma_\delta = 0.077$ and $\zeta = -6.45$. For each sample, 1 to 4, the top figure shows the generated data, y , and the bottom figure shows the generated mismatch error, δ .

The simulated data in Fig. 19 looks much closer to the measured data in Fig. 16, however, our MAP estimate of ζ has now significantly changed, we ought to see if this has had an effect on the 3rd parameter, σ_ϵ . To do this we will fix $\zeta = -6.45$ and sample the posterior over σ_ϵ and σ_δ , as shown in Fig. 20.

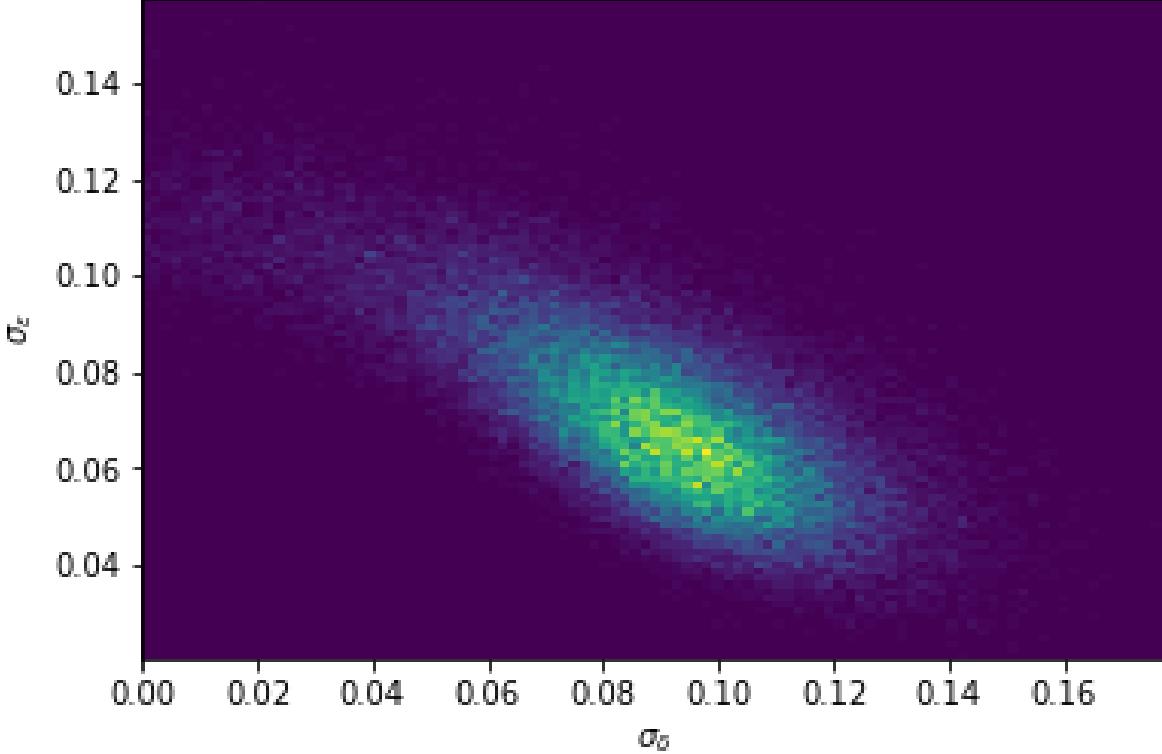


Figure 20: Histogram showing dependence of σ_δ and σ_ϵ . 45,011 samples with 65.3% acceptance rate.

Fig. 20 shows there is some dependence on σ_δ and σ_ϵ . This is because if less of the total error, $\boldsymbol{\delta} + \boldsymbol{\epsilon}$, can be explained by $\boldsymbol{\epsilon}$, then more of the error has to be explained by $\boldsymbol{\delta}$. Therefore the scale value of the covariance matrix \mathbf{K} , σ_δ , must be larger, and vice versa so there is a negative correlation between values of σ_ϵ and σ_δ .

This suggests a more complex relationship between the model parameters, ζ , σ_ϵ and σ_δ . Further work should be done to generate the 3-dimensional posterior over these parameters.

6.1 Future work

As aforementioned, further work should be done to generate the 3-dimensional posterior over the model parameters, ζ , σ_ϵ and σ_δ .

There are a few more things that we haven't yet considered. The mismatch errors shown in this report have tended to be underestimates, or overestimates, of the deflections in the models. This may be better modelled by a scaling coefficient, ρ , on the model deflection values, \mathbf{u}_h . Furthermore, the sensor locations may not be in the same locations as the node positions of the FEM model. In this case a projection matrix, Φ , must be used to project the model deflection values, \mathbf{u}_h , on to the sensor locations. The random mismatch error then becomes,

$$\boldsymbol{\eta} = \rho \Phi \mathbf{u}_h + \boldsymbol{\delta}, \quad (28)$$

with,

$$\boldsymbol{\delta} \approx N(0, \sigma_\delta^2 \mathbf{K}). \quad (29)$$

Further work could be done to infer the posteriors over these parameters, and investigate the dependence between the four model parameters, σ_δ , σ_ϵ , λ (or ζ) and ρ .

6.2 Practical considerations

If sensors are to be used in infrastructure, expense will be proportional to the number of sensors installed, rather than the number of readings taken. It is therefore a good thing that this report has shown that statistical parameters can be reasonably well inferred by just 8 sensors, by taking multiple readings.

The beam had developed a crack impacted the deflection data (see Section. 5.4). This outlines a possible advantage of using sensors in infrastructure: by monitoring the inferred model parameters over time, we may be able to detect changes in these parameters that indicate fatigue or other damage.

7 Conclusions

- Deflection data of the gyroid beam, a shell structure, can be modelled using Bayesian parameters.
- The data was successfully modelled using two covariance matrix scaling parameters, σ_δ , σ_ϵ and a length scale parameter, λ . Further work needs to be done to investigate if the model is better explained with a fourth parameter, a scaling coefficient, ρ . Further work needs to be done to determine the dependence between these parameters.
- The model parameters, σ_δ , σ_ϵ and λ , were successfully inferred using just 8 sensors.
- The assumption that observation error, ϵ , along the beam is IID (independent identically distributed) is not always correct, and will depend on the measuring apparatus.

8 Appendix

8.1 Normally distributed data from a Bayesian point of view, and calculating the marginal likelihood

This section is adapted from the lecture notes "The Conjugate Prior for the Normal Distribution", written by Michael I. Jordan [9].

We will look at the Gaussian distribution from a Bayesian point of view. It is first assumed that the deflection data is normally distributed about a known mean,

$$y_i \sim \mathcal{N}(\mu, \sigma^2), \quad (30)$$

where,

$$\mu = u(x_i; \theta) = u_i. \quad (31)$$

Given that u is fixed, then the conjugate prior (if this term is unfamiliar, then read through chapters 1-4 of Simon Rogers, S. & Girolami, M. 'A first course in machine learning' Chapman & Hall) for σ^2 is an inverse gamma distribution,

$$\sigma^2 \sim IG(\alpha, \beta). \quad (32)$$

If we re-parametrize in terms of precisions, the conjugate prior is a Gamma distribution,

$$\tau \sim Ga(\alpha, \beta), \quad P(\tau|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \tau^{\alpha-1} e^{-\tau\beta}. \quad (33)$$

We want to find the marginal likelihood of our data given the model parameters, α and β . This can be done by marginalizing out τ ,

$$P(y|u, \alpha, \beta) = \int P(y|u, \tau) P(\tau|\alpha, \beta) d\tau. \quad (34)$$

This integral "smears" the Gaussian into a heavier tailed distribution, which will turn out to be the student's t-distribution,

$$\begin{aligned} P(y|u, \alpha, \beta) &= \int \frac{\beta^\alpha}{\Gamma(\alpha)} \tau^{\alpha-1} e^{-\tau\beta} \left(\frac{\tau}{2\pi} \right)^{\frac{1}{2}} \exp \left(-\frac{\tau}{2}(y-u)^2 \right) d\tau \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{1}{\sqrt{2\pi}} \int \tau^{(\alpha+\frac{1}{2})-1} e^{-\tau(\beta+(y-u)^2)/2} d\tau \quad \text{Gamma integral; use normalising constant} \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{1}{\sqrt{2\pi}} \frac{\Gamma(\alpha + \frac{1}{2})}{(\beta + \frac{1}{2}(y-u)^2)^{\alpha+\frac{1}{2}}} \\ &= \frac{\Gamma(\alpha + \frac{1}{2})}{\Gamma(\alpha)} \frac{1}{(2\pi\beta)^{\frac{1}{2}}} \frac{1}{(1 + \frac{1}{2\beta}(y-u)^2)^{\alpha+\frac{1}{2}}}. \end{aligned} \quad (35)$$

It follows that $\sqrt{\frac{\alpha}{\beta}}(y-u)$ follows a student's t-distribution with 2α degrees of freedom.

Given the assumption that the noise, $\varepsilon_i = y_i - u_i$, is IID, the joint conditional density can be factorised into N separate terms, one for each data object, and the marginal likelihood, L , of the dataset is in product form,

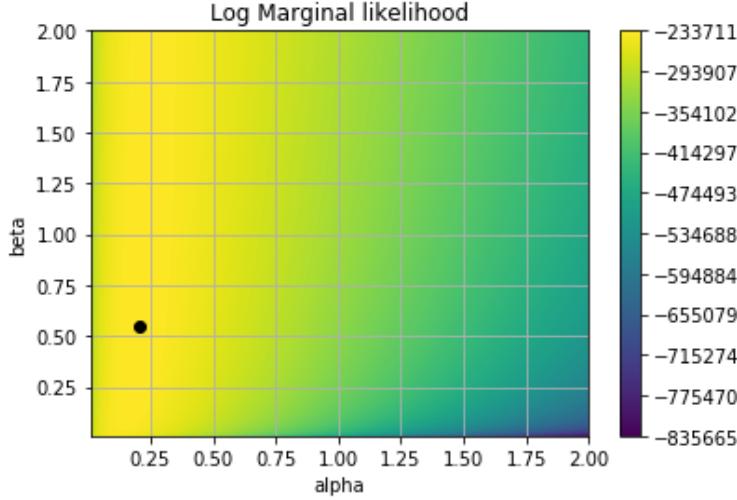
$$L = P(\mathbf{y}|\mathbf{u}, \alpha, \beta) = \prod_i P(y_i|u_i, \alpha, \beta). \quad (36)$$

Note that we haven't gone as far as saying that the y_i values are themselves completely independent. They are conditionally independent - given a value for u_i , the deterministic part of the model, the y_i are independent, meaning there should be no pattern in the error.

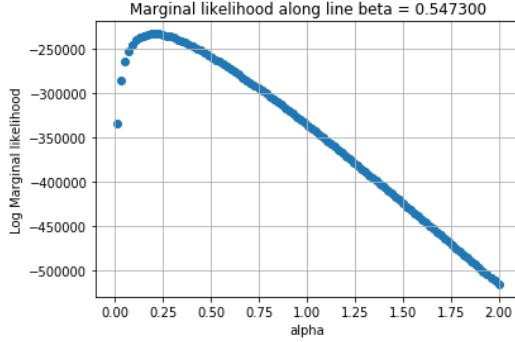
8.1.1 An illustrative example with example data

We generate some data from the student's t-distribution, \mathbf{y} around a mean, \mathbf{u} , using prescribed hyper-parameters $\alpha = 0.2$ and $\beta = 0.5$. Suppose we didn't know the underlying parameters, α and β that generated the data. We calculate the marginal likelihood of the data to optimise α and β . This is known as the maximum likelihood (ML) estimate of α and β . Assuming that α and β can take any value in the ranges

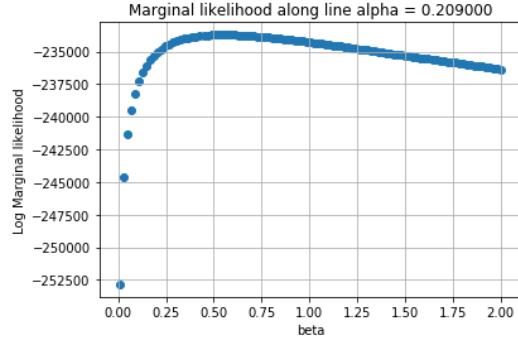
$$\begin{aligned} 0 \leq \alpha &\leq 2 \\ 0 \leq \beta &\leq 2 \end{aligned}$$



(a) Marginal likelihood contour plot as a function of the prior parameters, α and β . The circle on the left shows the maximum likelihood estimate.



(b) Log marginal likelihood as a function of α , with β optimum.



(c) Log marginal likelihood as a function of β , with α optimum.

Figure 21: Marginal likelihood plots showing the likelihood of the data given the Euler-Bernoulli beam model and parameters α and β .

Figure 21a shows the log marginal likelihood as α and β are varied in their respective ranges. The ML estimate $\alpha = 0.209$ and $\beta = 0.547$ lies close to our prescribed values of $\alpha = 0.2$ and $\beta = 0.5$ that were used to generate the data \mathbf{y} .

8.2 Monte Carlo Procedures

8.2.1 An illustrative example with example data

To test the model, and to see if the proposed method works, we will generate 'good' data using prescribed values of $\sigma_\epsilon = 0.6$ and $\lambda = 0.00823$ and run the Markov chain using the data to obtain a posterior over σ_ϵ and λ . Generating 'good' data is necessary to test the proposed method. To make things simple, model displacements, errors and measured displacements have all been generated at the same 'sensor locations'. For speed, instead of using a Finite Element package, the Euler-Bernoulli beam model was used to generate the model data, \mathbf{u}_h . This example data is shown in Fig. ??.

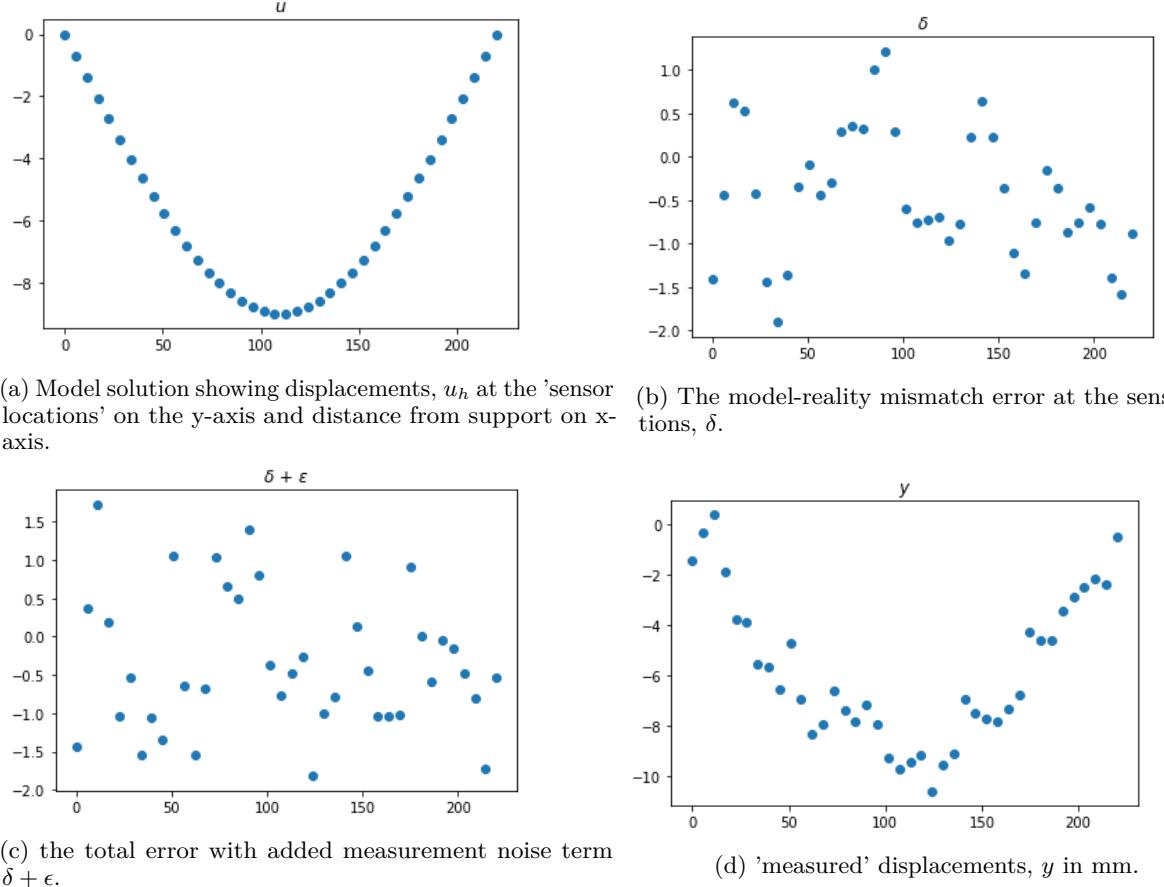


Figure 22: Example data.

For a first try, the proposal density covariance was chosen to be $\Sigma = \begin{bmatrix} 0.2 & 0 \\ 0 & 0.2 \end{bmatrix}$ and the first sample location was chosen to be $\sigma_\epsilon = 1.0$, $\lambda = 1.0$.

The Metropolis algorithm was ran for a long period of time (100,000 iterations, taking 500s). The acceptance rate (after counting burn in samples as rejected) was low at 4.007%, giving 4007 samples. A burn was performed on the first 200 samples: this was enough samples for the Markov chain to converge on the distribution.

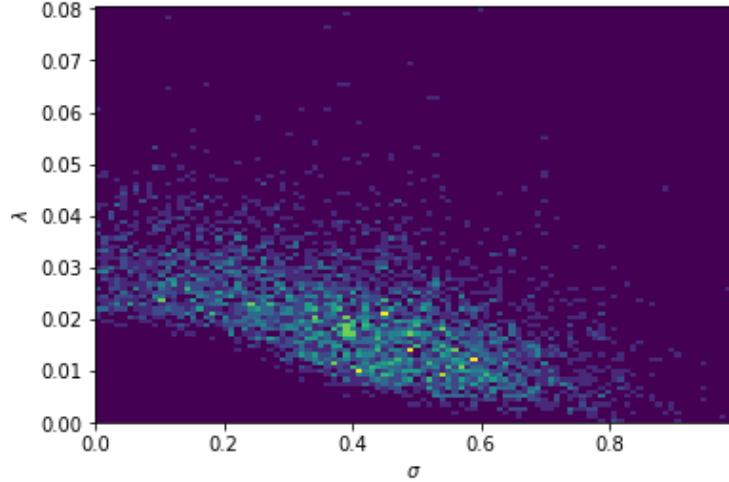


Figure 23: histogram of 4007 samples of lambda, sigma pairs

The histogram seems to show convergence, given that the start point was chosen at $\sigma_{\epsilon,0} = 1.0$, $\lambda_0 = 1.0$. The posterior density picks out a high likelihood at the prescribed value of $\sigma_\epsilon = 0.6$ and $\lambda = 0.00823$, however, there is a sharp drop off of samples close to $\lambda = 0$. This is because if the proposal sample is negative or 0, then it is always rejected (given the gamma prior), so there will be a sharp drop off of samples close to 0, skewing the distribution. To solve this issue, the proposal density was made 200 times smaller in each direction, and a closer starter location to lambda was picked $\lambda_0 = 0.1$ so that the chain would converge quicker.

Another point to note is the shape of the distribution. The dependence of λ on σ_ϵ may be because, for example, in the hypothetical case that none of the error can be explained by σ_ϵ (i.e. say $\sigma_\epsilon = 0$), then more of the error has to be explained by λ . Therefore the covariance matrix K , generated by λ must look more like IID, so λ must get larger, increasing the length scale above the dimension of the beam.

As well as decreasing the size of the proposal distribution $\Sigma = \begin{bmatrix} 0.001 & 0 \\ 0 & 0.001 \end{bmatrix}$, the number of iterations was doubled to 200,000. As expected, the acceptance rate was higher at a reasonable 19.778% giving 39,556 samples. The histogram looks much the same as before with a high, but by no means highest, amount of samples close to the 'true' values of $\sigma_\epsilon = 0.6$ and $\lambda = 0.00823$

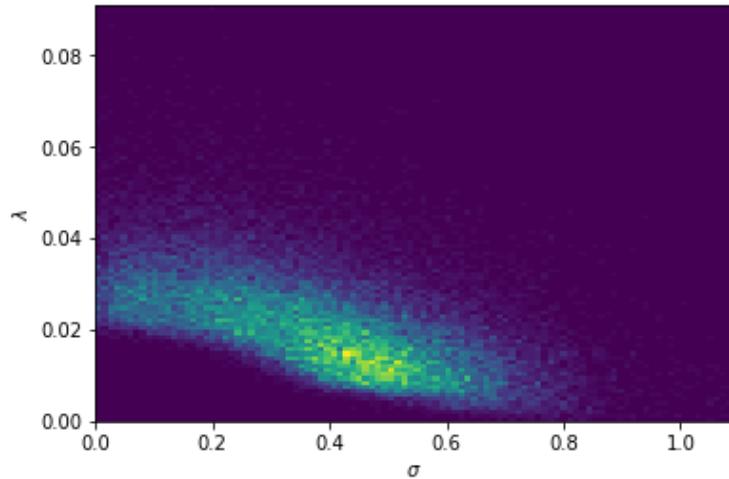
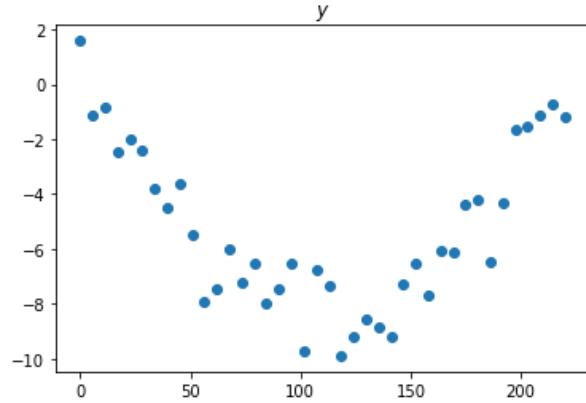
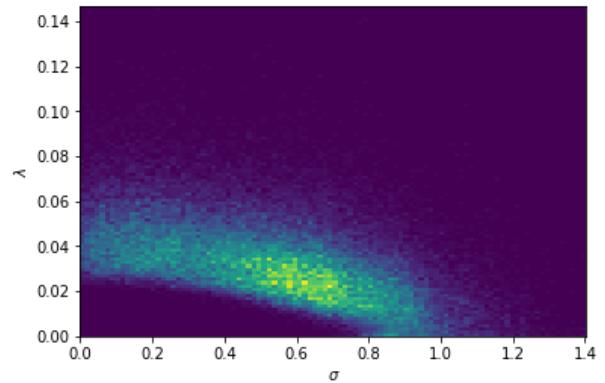


Figure 24: histogram of 39556 samples of λ , σ_ϵ pairs

The data used to produce these distributions is convenient. Let us try using different values of lambda and sigma with the same proposal densities and priors.

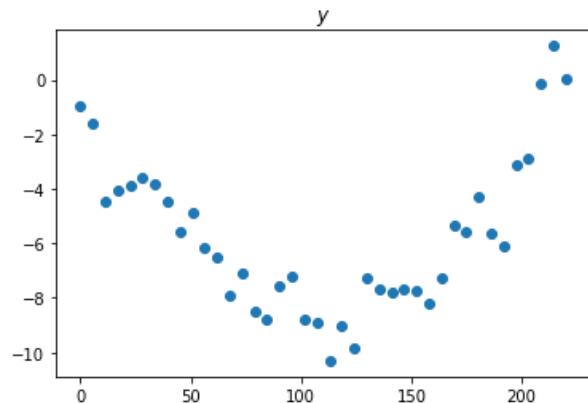


(a) Measured displacements in mm.

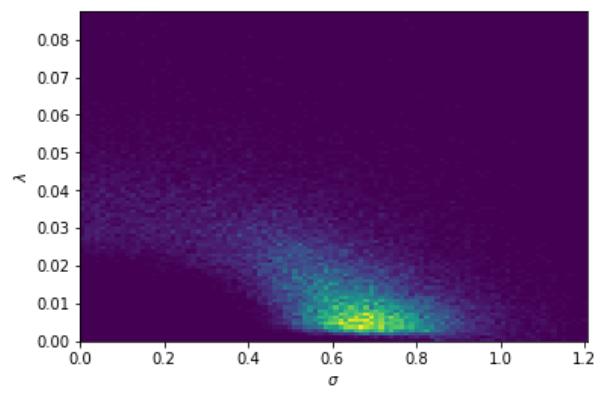


(b) Histogram of samples.

Figure 25: Samples generated with values $\sigma_\epsilon = 0.6$, $\lambda = 0.04$

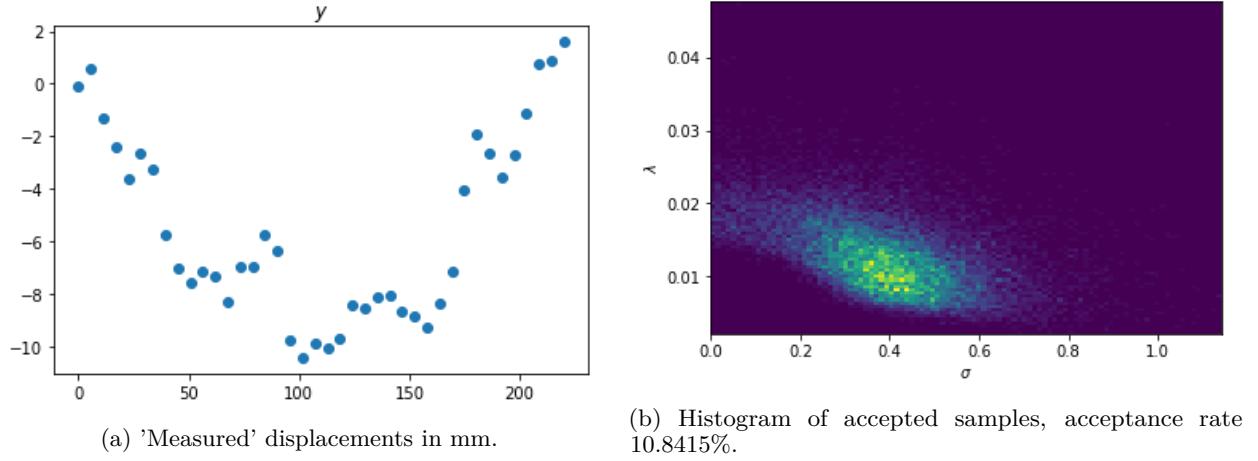
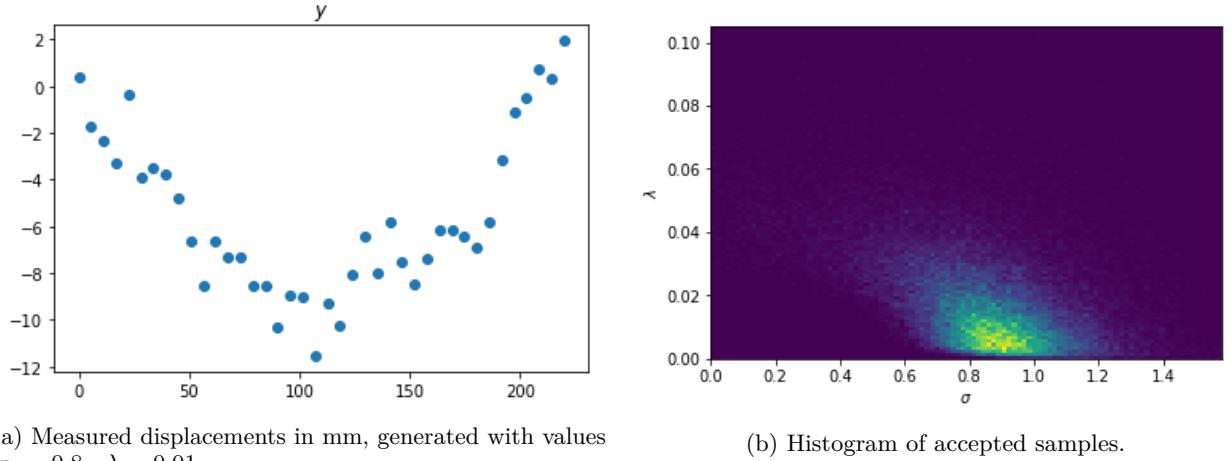


(a) Measured displacements in mm.



(b) Histogram of samples.

Figure 26: Samples generated with values $\sigma_\epsilon = 0.4$, $\lambda = 0.04$.

Figure 27: Samples generated with values $\sigma = 0.4$, $\lambda = 0.01$.Figure 28: Samples generated with values $\sigma_\epsilon = 0.8$, $\lambda = 0.01$.

8.2.2 Changing variables

The low sample rate is likely to be a problem with how close to 0 the lambda is. The likelihood of samples close to zero is very low since if the proposal sample is negative or 0, then it is always rejected, so likelihood will have a sharp drop off close to 0, skewing the distribution. To solve this issue, we will attempt make a change of variables

$$\lambda = \exp(\zeta) \quad (37)$$

For a continuous random variable, the relationship between the cdf

$$F_x(x) = P(X \leq x) = \int_{-\infty}^x f_x(t) dt \quad (38)$$

and its pdf $f_x(t)$ is

$$f_x(t) = \frac{dF_x(t)}{dx} \quad (39)$$

The cdf is useful when characterising the probability distribution of a transformation of a random variable. That is from X define new random variables, e.g. $Y = \exp(X)$ or $Y = X^2$.

Let X have pdf f_X and cdf F_X . Define the new random variable $Y = r(X)$ and then follow these steps to derive the pdf f_Y .

1. For each y find the set $A_y = \{x : r(x) \leq y\}$

2. Find the cdf of Y .

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(r(X) \leq y) \\ &= P(X \in A_y) \\ &= \int_{-\infty}^{\infty} \mathbb{1}_{A_y}(x) f_x(x) dx \end{aligned} \tag{40}$$

3. Let $F'_y(y)$ be the derivative of $F_y(y)$ at y (when it exists). Set $f_Y(y) = F'_y(y)$

When r is strictly increasing or strictly decreasing we can derive a formula for f_Y . In this case r has an inverse, let $s = r^{-1}$. Then

$$f_Y(y) = f_X(s(y)) \left| \frac{ds(y)}{dy} \right| \tag{41}$$

Making the suggested substitution, $\zeta = \ln \lambda$, such that

$$s(\zeta) = \exp(\zeta) \tag{42}$$

since $r(\lambda) = \ln |\lambda|$ is strictly increasing, then we can derive a formula for f_ζ .

$$\begin{aligned} f_\zeta(\zeta) &= f_\lambda(s(\zeta)) \left| \frac{ds(\zeta)}{d\zeta} \right| \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \exp(\zeta)^{\alpha-1} \exp(-\beta \exp(\zeta)) \left| \frac{d(\exp(\zeta))}{d\zeta} \right| \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \exp(\zeta(\alpha-1) - \beta \exp(\zeta) + 1) \\ &= \frac{\beta^\alpha}{\Gamma(\alpha)} \exp(\zeta\alpha - \beta \exp(\zeta)) \end{aligned} \tag{43}$$

This distribution, with $\alpha = 1.0$ and $\beta = 100.0$ is shown in Fig. 29a. This distribution, with $\alpha = 1.0$ and $\beta = 300.0$ is shown in Fig. 29b.

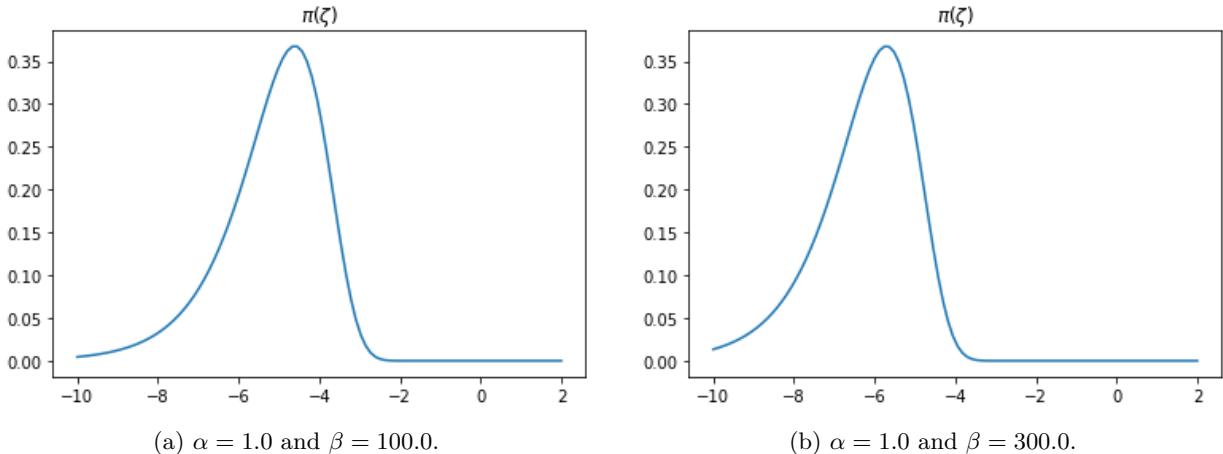
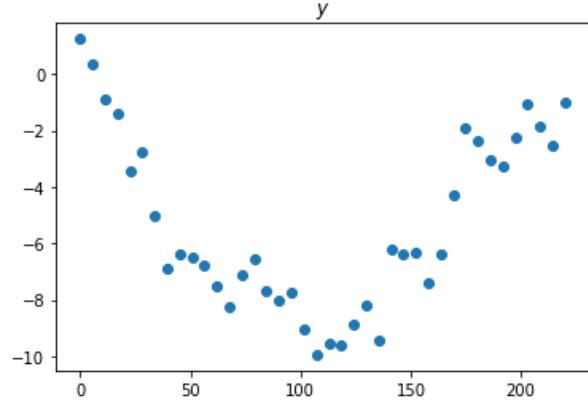
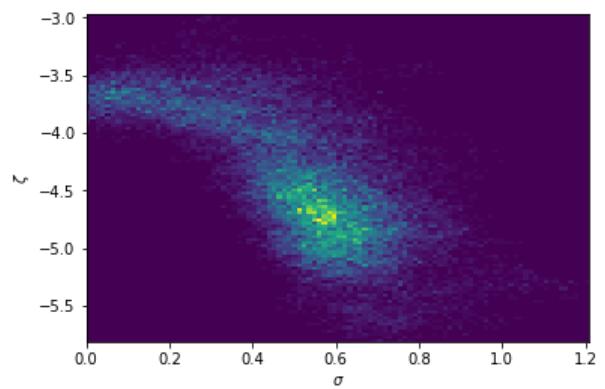


Figure 29: Prior on ζ

The result of this change of variables drastically increases the acceptance rate of our sampler to between 80% – 90%, which is a good thing. Generating data with values of $\lambda = 0.00823$ (i.e. $\zeta = -4.8$) and $\sigma_\epsilon = 0.6$ and the prior shown in Fig. 29a gives the following results.

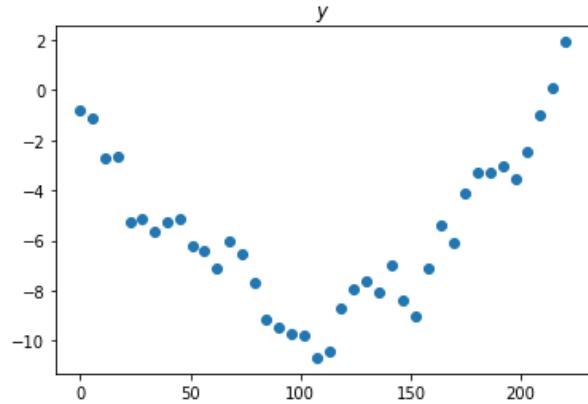


(a) Measured displacements in mm, generated with values $\sigma_\epsilon = 0.8$, $\zeta = -4.8$.

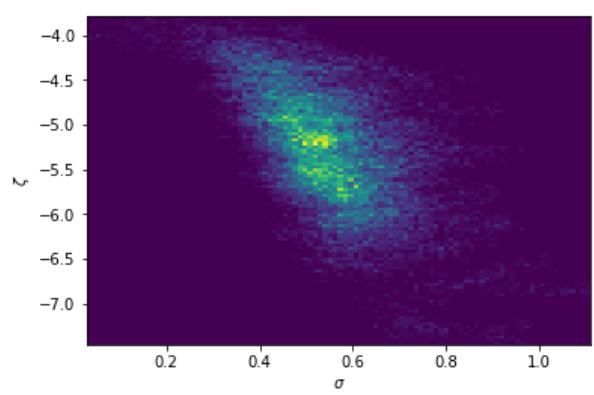


(b) Histogram of accepted samples, acceptance rate 86.9%, using the prior shown in Fig.29a.

with the same data and a different prior, shown in Fig.29b gives the following results



(a) Measured displacements in mm, generated with values $\sigma_\epsilon = 0.8$, $\zeta = -4.8$.



(b) Histogram of accepted samples, using the prior shown in Fig.29b.

This shows that the target distribution is strongly dependent on the parameters put on the prior. Therefore, the prior should be selected carefully.

8.2.3 Multiple sets of observations

Using multiple sets of sensor data will reduce the dependence of the posterior on the prior, and will result in a target distribution that has higher likelihood around the 'true' values of ζ and σ_ϵ .

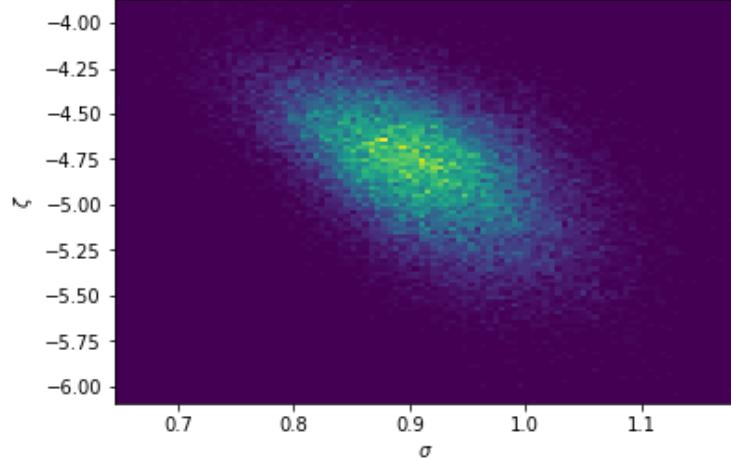


Figure 32: Histogram of accepted samples for $L = 8$ sets of observations, acceptance rate 49.1%, using the prior shown in Fig.29a. The posterior distribution has concentrated more around the 'true' values of $\sigma_\epsilon = 0.8$ and $\zeta = -4.8$.

Another result is that the distribution is much less dependent on the prior distribution.

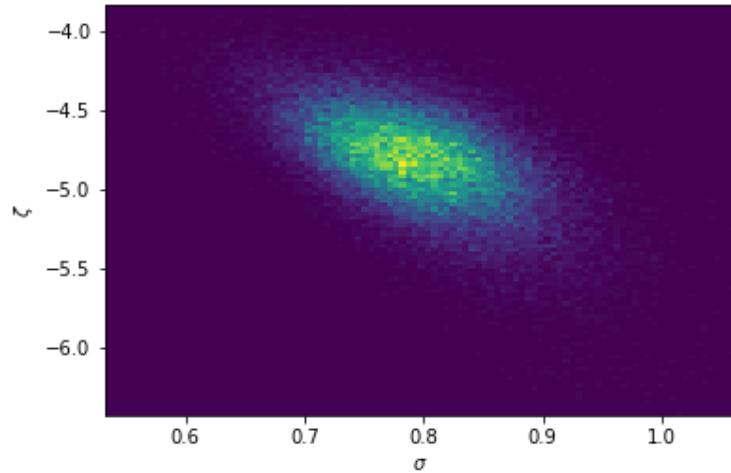


Figure 33: Histogram of 40,000 accepted samples for $L = 8$ sets of observations, acceptance rate 44.6%, using the prior shown in Fig.29b. The posterior distribution has concentrated more around the 'true' values of $\sigma_\epsilon = 0.8$ and $\zeta = -4.8$, just like in Fig.32, despite the fact that a different prior was put on ζ .

8.3 Raw data

All raw data can be viewed on the following google spreadsheet:
https://docs.google.com/spreadsheets/d/1JPFpRM5JMnOn2Zp7efNToLY_dl6NOjCAqsgbzGLFsLM/edit?usp=sharing

References

- [1] "A digital twin of the world's first 3D printed steel bridge" The Alan Turing Institute (2018) <https://www.turing.ac.uk/research/research-projects/digital-twin-worlds-first-3d-printed-steel-bridge> (accessed July 2019).
- [2] Hussein, A., Hao, L., Yan, C., Everson, R., and Young, P. (2013). Advanced lattice support structures for metal additive manufacturing. *Journal of Materials Processing Technology*, 213:1019–1026.
- [3] Abueidda, D. W., Bakir, M., Abu Al-Rub, R. K., Bergström, J. S., Sobh, N. A., and Jasiuk, I. (2017). Mechanical properties of 3d printed polymeric cellular materials with triply periodic minimal surface architectures. *Materials*
- [4] Fleck, N. A., Deshpande, V. S., and Ashby, M. F. (2010). Micro-architected materials: past, present and future. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 466:2495–2516.
- [5] Girolami, M., Gregory, A. Yin, G., Cirak, F. (2019). The Statistical Finite Element Method
- [6] Cirak, F. and Long, Q. (2011). Subdivision shells with exact boundary control and non-manifold geometry. *International Journal for Numerical Methods in Engineering*, 88:897-923.
- [7] Cirak, F. and Ortiz, M. (2001). Fully C^1 -conforming subdivision elements for finite deformation thin-shell analysis. *International Journal for Numerical Methods in Engineering*, 51:813-833.
- [8] Cirak, F., Ortiz, M., and Schröder, P. (2000). Subdivision surfaces: A new paradigm for thin-shell finite element analysis. *International Journal for Numerical Methods in Engineering*, 47:2039-2072.
- [9] "The Conjugate Prior for the Normal Distribution" Michael I. Jordan (2010). The Conjugate Prior for the Normal Distribution. Stat260: Bayesian Modelling and Inference, University of California, which can be viewed at <https://people.eecs.berkeley.edu/~jordan/courses/260-spring10/lectures/lecture5.pdf> (accessed July 2019).