

The Statistical Finite Element Method

Mark Girolami^{1,2}, Alastair Gregory^{2,3}, Ge Yin¹, and Fehmi Cirak¹

¹Department of Engineering, University of Cambridge, Cambridge

²Lloyd's Register Foundation's Programme for Data-Centric Engineering, Alan Turing
Institute, London

³ Department of Mathematics, Imperial College London, London

March 9, 2021

Abstract

The finite element method (FEM) is one of the great triumphs of modern day applied mathematics, numerical analysis and algorithm development. Engineering and the sciences benefit from the ability to simulate complex systems with FEM. At the same time the ability to obtain data by measurements from these complex systems, often through sensor networks, poses the question of how one systematically incorporates data into the FEM, consistently updating the finite element solution in the face of mathematical model misspecification with physical reality. This paper presents a statistical construction of FEM which goes beyond forward uncertainty propagation or solving inverse problems, and for the first time provides the means for the coherent synthesis of data and FEM.

Keywords: *Finite element models, Bayesian model updating, Model discrepancy*

1 Introduction

Mathematical models of physical and natural systems have long been employed to investigate the underlying mechanisms leading to the form and function of such processes being studied, with Partial Differential Equations (PDE) being almost ubiquitous in such scientific enquiry. The finite element method (FEM) (Strang and Fix, 1973) has emerged as the dominant approach to numerically approximate the solutions of PDEs describing natural and physical systems. The FEM has revolutionised the engineering, physical and life sciences by enabling large scale computer simulations that are integral to the study of the physical systems they model.

These mathematical models and associated computer simulations are in many cases simplifications of the actual system, for example making assumptions about the nonlinearity of certain system processes or model parameters, thus introducing possible model misspecification (Kennedy and O'Hagan, 2001). In addition to this, computer simulations of these models are often extremely computationally expensive to run. It is because of this that such models are seldom used as the only component in the overall modelling procedure; data observed from the actual physical system, obtained as measurements from sensors, is often utilised as well. The mathematical models can be calibrated, tuned and combined with this observed data. In most cases, this data is used to infer information about the undefined parameters in the model (Stuart, 2010; Grafe, 1998). Furthermore uncertainty in the parameters or forcing, of the possibly misspecified model, is captured and forward propagated in such computer simulations, see for example Ghanem and Spanos (2003).

The connection between sensor data and the FEM is currently restricted to data assimilation for solving inverse problems or the calibration of PDE based models (Stuart, 2010). This however places unwarranted faith in the fidelity of the underlying mathematical description of the actual system under study (Gregory et al., 2019). As originally argued in Kennedy and O'Hagan (2001), if one concedes

that there is missing physics, chemistry or some such misspecification between generative reality and the mathematical abstraction, defining the FE model, then a framework to systematically characterise and propagate this uncertainty in the FE model is required.

This paper introduces a novel unifying approach to provide a fully statistical FEM where both the FE model and observational data are integrated into a coherent inferential framework. This provides the means to fuse data and model where model misspecification is a reality and observational data can be employed to provide data adjusted FEM solutions. This addresses some of the shortcomings of FEM in an era where sensor measurements and data are becoming widely and inexpensively available in scientific enquiry and technological development. The framework is constructed by the introduction of a randomised discrepancy term representing unknown model-reality mismatch or low fidelity representation (Kennedy and O’Hagan, 2001), in the design of a generative data-model driven from the FE model. The next section will first introduce a probabilistic representation of an FE model, before Sec. 3 discusses the misspecification of such models to observed data. Then a theoretical and numerical presentation of the proposed framework follows in Sec. 4 and Sec. 5.

2 Probabilistic Representation of the FE Method

Consider a linear differential operator equation which generalise a large class of Partial Differential Equations describing many physical processes in e.g. thermodynamics and continuum mechanics. The case for nonlinear operator equations will be discussed at the end of the paper. We assume that the mathematical representation of the process, e.g. the deflection of a bridge under loading, or the temperature distribution in a jet engine, is modelled by $u(x)$. This is governed by the partial differential equation $\mathcal{L}_\theta u(x) = -f(x)$, where $\mathcal{L}_\theta$ is some randomised self-adjoint elliptic linear operator over space, with respect to the uncertain parameters θ (e.g. material properties such as thermal conductivity or bulk material stiffness). The governing equation is to be equipped with suitable boundary conditions which are omitted for the sake of brevity. Furthermore the uncertainty induced due to incomplete knowledge of the right hand side forcing term $f(x)$ will be formally incorporated by defining a probabilistic model of the randomised forcing.

For the purposes of developing ideas let us assume that the statistical characteristics of the right hand side error are represented by an appropriately defined additive infinite dimensional Gaussian measure, having covariance function C_u with some lengthscale parameter σ_f ,

$$\mathcal{L}_\theta u(x) = -f(x) - \epsilon_f(x), \quad \epsilon_f(x) \sim \mathcal{GP}(0, C_f).$$

The push-forward of this measure induces a conditional infinite dimensional Gaussian measure over the Hilbert space of solutions $u(x)$ such that,

$$p(u(x)|f(x), \theta) = \mathcal{GP}(-\mathcal{L}_\theta^{-1}f(x), \mathcal{L}_\theta^{-1}C_f\mathcal{L}_\theta^{-1}),$$

with $\mathcal{L}_\theta^{-1}$ the Green’s function for the PDE. Now 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 equivalent randomised weak form,

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

where $\mathcal{B}_\theta$ is the randomised bilinear form corresponding to $\mathcal{L}_\theta$ and $\langle \cdot, \cdot \rangle$ is the appropriate Hilbert space inner-product. In standard finite element (FE) theory 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 I with a countable basis index set I_h . For FE models one obtains, $\sum_{i,j \in I_h} u_i \mathcal{B}_\theta(\phi_i(x), \phi_j(x))$ is equal to $\langle f(x), \phi_j(x) \rangle + \langle \epsilon_f(x), \phi_j(x) \rangle$. Finally the finite-dimensional conditional Gaussian measure over the FE basis coefficients is obtained,

$$\begin{aligned} p(\mathbf{u}_h | \mathbf{A}, f, \theta, h) &\equiv p(u_h(\mathbf{x}_u) | \mathbf{A}, f, \theta, h) \\ &= \mathcal{N}(\mathbf{A}^{-1} \mathbf{b}, \mathbf{A}^{-1} \mathbf{G} \mathbf{A}^{-1}) \\ &\equiv \mathcal{N}(\mathbf{m}_u, \mathbf{C}_u), \end{aligned} \tag{1}$$

where the elements of the matrices above are each $A_{i,j} = \mathcal{B}_\theta(\phi_i(x), \phi_j(x))$, $b_j = \langle f(x), \phi_j(x) \rangle$, $G_{i,j} = \langle \phi_i(x), C_f \phi_j(x) \rangle$ and $\mathbf{u}_h = u_h(\mathbf{x}_u)$ is our randomised FE approximation to the function $u(x)$ at the spatial coordinates $\mathbf{x}_u$.

Taking a closer look at this probabilistic construction, the first thing to notice is that the mean of the N -dimensional multivariate Gaussian $\mathbf{m}_u = \mathbf{A}^{-1}\mathbf{b}$ is exactly the Galerkin projection solution for the FEM (Strang and Fix, 1973). The covariance structure, $\mathbf{C}_u = \mathbf{A}^{-1}\mathbf{G}\mathbf{A}^{-1}$, corresponds to the action of the linear PDE operator combined with the projected right hand side noise onto the FE basis with $\mathbf{G}$, a Galerkin discretized covariance function. As the variance of the right hand side noise collapses to zero i.e. there is no uncertainty associated with the forcing term in the PDE, then we see that $p(\mathbf{u}_h|\mathbf{A}, f, \theta, h) = \delta(\mathbf{A}^{-1}\mathbf{b})$ is the conditional Dirac measure around the Galerkin solution. In other words we recover the deterministic Galerkin solution. This is of course all in the absence of data obtained from physical measurements. The following section now considers how the probabilistic FEM solution can be updated with the availability of data.

3 Finite Element Models Conditioned on Data

Consider a vector of operational data, denoted as $\mathbf{y} \in \mathbb{R}^{N \times 1}$. This N dimensional vector may represent N measurements taken simultaneously from N sensors, or in a time evolving manner. At present we make the reasonable working assumption that this data is acquired with independent Gaussian measurement error with scale parameter σ_ϵ . This of course can be generalised to capture more complex error structures. The important point to note here is that a statistical model of the data posits a *true* underlying generating process, which could be described if a full knowledge of the physical process was available. As such the observed measured data is a linear combination of the measurement error and the response of the true unknown generating process. Note that the mathematical model, the FE model in this case, will be a, possibly highly, misspecified representation of the true underlying generating process. Define $\Phi \in \mathbb{R}^{N \times M}$ to be the interpolation matrix of basis function evaluations that transforms the domain of the FEM solution $\mathbf{u}_h \in \mathbb{R}^{M \times 1}$, at the points $\mathbf{x}_u$, to the domain of the data $\mathbf{y}$, at the points $\mathbf{x}_y$ (i.e. where the sensors are located physically).

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 $\boldsymbol{\eta} \in \mathbb{R}^N$. As $\boldsymbol{\eta}$ is unobserved its value is considered as a random function, and one approach to model this random value is to incorporate this additional layer of uncertainty in the statistical model. One therefore posits that the randomised values of predictions from the FE model, denoted as $\Phi\mathbf{u}_h$, forms an additive regression model with scaling coefficient ρ and random mismatch error functional $\boldsymbol{\delta} \in \mathbb{R}^N$ such that $\boldsymbol{\eta} = \rho\Phi\mathbf{u}_h + \boldsymbol{\delta}$ with $\boldsymbol{\delta} \sim \mathcal{N}(\mathbf{0}, \mathbf{K})$. Here, $\mathbf{K} \in \mathbb{R}^{N \times N}$ is a matrix of covariance function evaluations $k_\delta(x, x')$, with lengthscale l , evaluated at the points $\mathbf{x}_y$.

The final component in the data model is the observation 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 $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma_\epsilon^2 \mathbf{I})$. The conditional probability of the data under random vectors drawn from both the FE model probability measure and the measure of the FE model-data mismatch, gives $p(\mathbf{y}|\mathbf{u}_h, \mathbf{A}, f, h, \theta, \sigma_f, \boldsymbol{\delta}, l, \rho, \sigma_\epsilon) = \mathcal{N}(\rho\Phi\mathbf{u}_h + \boldsymbol{\delta}, \mathbf{I}\sigma_\epsilon^2)$, where ρ is a regression term and $\mathbf{I} \in \mathbb{R}^{N \times N}$ is the identity matrix. The marginal probability of the data, $p(\mathbf{y}|\mathbf{A}, f, h, \theta, \sigma_f, l, \rho, \sigma_\epsilon)$, where the random vectors $\mathbf{u}_h$ and $\boldsymbol{\delta}$ can be integrated out, explicitly in this case¹, follows as,

$$\mathcal{N}(\rho\Phi\mathbf{m}_u, \rho^2\Phi\mathbf{C}_u\Phi^T + \mathbf{K} + \mathbf{I}\sigma_\epsilon^2). \quad (2)$$

Let us take a moment to consider this finite dimensional probability distribution to obtain some insight regarding the model and the assumptions being made. Firstly the measured data is multivariate Gaussian whose mean is defined by the ρ -scaled interpolation of the FEM solution $\rho\Phi\mathbf{m}_u$. One can observe that it is being explicitly assumed that the FE model, once scaled, is providing a component of the underlying mean of the data $\rho\Phi\mathbf{m}_u$. Recall that $\mathbf{m}_u = \mathbf{A}^{-1}\mathbf{b}$ is the mean of the Gaussian probability measure associated with the FE model, and this is exactly the Galerkin solution of the FEM. We therefore observe that from

¹Under different model and distributional assumptions Monte Carlo procedures may be required to numerically perform this marginalisation.

the probabilistic description, existing numerical methods (FEM in this case) appear in the theoretically principled systematic probabilistic construction.

4 Inference in the Statistical FE Method

Now that the probabilistic representation has been defined, we can proceed in making statistical inference based on data and the underlying FE model. Recall that the vector $\mathbf{u}_h$ of FE coefficients has, prior to making observations from the physical system, a Gaussian measure given by,

$$p(\mathbf{u}_h | \mathbf{A}, f, h, \boldsymbol{\theta}) = \mathcal{N}(\mathbf{A}^{-1}\mathbf{b}, \mathbf{A}^{-1}\mathbf{G}\mathbf{A}^{-1}) \equiv \mathcal{N}(\mathbf{m}_u, \mathbf{C}_u). \quad (3)$$

It is interesting to note that the covariance structure, and hence statistical error structure, is defined by the conditioning of the FE stiffness matrix based on the spectral distribution of the actual differential operator $\mathcal{L}_\theta$. For clarity of presentation, the conditioning of the probability distributions on the set of variables $\{\mathbf{A}, f, h, \boldsymbol{\theta}, \sigma_f, l, \rho, \sigma_\epsilon\}$ shall now be made implicit. Once measurements have been obtained from sensors and data $\mathbf{y}$ is now available from the physical process, it is straightforward from conditioning of Gaussian distributions to see that the distribution of the FE coefficients, conditional on the measured data, is now also Gaussian². Therefore the probabilistic construction of the statistical FE model has the following components,

$$p(\mathbf{u}_h) = \mathcal{N}(\mathbf{m}_u, \mathbf{C}_u) \quad (4)$$

$$p(\mathbf{y} | \mathbf{u}_h) = \mathcal{N}(\rho \Phi \mathbf{u}_h, \mathbf{K} + \mathbf{I}\sigma_\epsilon^2) \quad (5)$$

$$p(\mathbf{y}) = \mathcal{N}(\rho \Phi \mathbf{m}_u, \rho^2 \Phi \mathbf{C}_u \Phi^\top + \mathbf{K} + \mathbf{I}\sigma_\epsilon^2) \quad (6)$$

$$p(\mathbf{u}_h | \mathbf{y}) = \mathcal{N}(\mathbf{m}_{u|y}, \mathbf{C}_{u|y}), \quad (7)$$

where

$$\mathbf{m}_{u|y} = \mathbf{C}_{u|y}(\rho^2 \Phi^\top (\mathbf{K} + \sigma_\epsilon^2 \mathbf{I})^{-1} (\mathbf{y}/\rho) + \mathbf{C}_u^{-1} \mathbf{m}_u) \quad (8)$$

$$\mathbf{C}_{u|y} = (\mathbf{C}_u^{-1} + \rho^2 \Phi^\top (\mathbf{K} + \sigma_\epsilon^2 \mathbf{I})^{-1} \Phi)^{-1}. \quad (9)$$

This provides a powerful systematic methodology in which to statistically update the Galerkin FEM solution $\mathbf{A}^{-1}\mathbf{b}$ when data from sensors is available and the mathematical model is misspecified. The marginal likelihood, probability of the data, $p(\mathbf{y})$ can be used alongside Markov chain Monte Carlo or optimisation routines such as simulated annealing to obtain statistical estimates of the model parameters and other unknowns e.g. $\{\sigma_f, l, \rho, \sigma_\epsilon\}$.

4.1 Computational Cost of Inference

By exploiting the Woodbury matrix identity, denoting in shorthand $\mathbf{K} + \sigma_\epsilon^2 \mathbf{I}$ as $\mathbf{K}_s$ and without loss of generality setting $\rho = 1$, the conditional mean $\mathbf{m}_{u|y}$ takes the form

$$[\mathbf{I} - \mathbf{C}_u \Phi^\top (\mathbf{K}_s + \Phi \mathbf{C}_u \Phi^\top)^{-1} \Phi] [\mathbf{C}_u \Phi^\top \mathbf{K}_s^{-1} \mathbf{y} + \mathbf{m}_u]. \quad (10)$$

In terms of the number of solves of a potentially large sparse $M \times M$ dimensional linear system of the form $\mathbf{A}\mathbf{x} = \mathbf{z}$ required to obtain $\mathbf{m}_{u|y}$, where iterative Krylov methods may need to be resorted to due to limited compute budget, it can be seen that an additional $2(N+1)$ i.e. $\mathcal{O}(N)$ Krylov based Galerkin solves are required, where N is the number of distinct sensors.

4.2 Posterior of the Generating Process

The posterior data-conditional for the posited true unobserved process $\boldsymbol{\eta}$ follows given that,

$$p(\boldsymbol{\eta}) = \mathcal{N}(\rho \Phi \mathbf{m}_u, \rho^2 \Phi \mathbf{C}_u \Phi^\top + \mathbf{K}) \equiv \mathcal{N}(\mathbf{m}_\eta, \mathbf{C}_\eta) \quad (11)$$

$$p(\mathbf{y} | \boldsymbol{\eta}) = \mathcal{N}(\boldsymbol{\eta}, \mathbf{I}\sigma_\epsilon^2) \quad (12)$$

$$p(\boldsymbol{\eta} | \mathbf{y}) = \mathcal{N}(\mathbf{m}_{\eta|y}, \mathbf{C}_{\eta|y}), \quad (13)$$

²where now the conditioning is implicit then $p(\mathbf{u}_h | \mathbf{y})$ should be understood as $p(\mathbf{u}_h | \mathbf{y}, \mathbf{A}, f, h, \boldsymbol{\theta}, \sigma_f, l, \rho, \sigma_\epsilon)$

with,

$$\mathbf{m}_{\eta|\mathbf{y}} = \mathbf{C}_{\eta|\mathbf{y}}(\sigma_\epsilon^{-2}\mathbf{y} + \mathbf{C}_\eta^{-1}\mathbf{m}_\eta) \quad (14)$$

$$\mathbf{C}_{\eta|\mathbf{y}} = (\mathbf{C}_\eta^{-1} + \sigma_\epsilon^{-2}\mathbf{I})^{-1}. \quad (15)$$

The scaling of the solves required to obtain these conditional distributions is $\mathcal{O}(N)$, the number of sensor locations, which will be far smaller than the number of nodes employed in the FEM.

4.3 Predictive Distributions from the Statistical FE Method

The prediction of data values where there are no sensors and measurements have not been made can now be considered. Note that if it is required to predict the values of data which would be measured at locations where instrumentation is unavailable, say $\mathbf{x}_n \in \mathbb{R}^d$, then the following distribution provides the means to do so using a hybrid of information from the FE model and data. Consider the data (unmeasured) $\mathbf{y}_n$ which would arise at coordinates $\mathbf{x}_n$. The distribution of possible values for this data is given by,

$$p(\mathbf{y}_n|\mathbf{y}) = \int p(\mathbf{y}_n|\mathbf{u}_h)p(\mathbf{u}_h|\mathbf{y})d\mathbf{u}_h,$$

noting that the matrix of FE interpolation bases can be written as $\Phi_n \in \mathbb{R}^{d \times M}$. Then we note that $p(\mathbf{y}_n|\mathbf{u}_h) = \mathcal{N}(\rho\Phi_n\mathbf{u}_h, \mathbf{K}_n + \mathbf{I}_n\sigma_\epsilon^2)$ with $\mathbf{K}_n \in \mathbb{R}^{d \times d}$ being the covariance function $\kappa_\delta(x, x')$ evaluated at $\mathbf{x}_n$ and $\mathbf{I}_n \in \mathbb{R}^{d \times d}$ being the identity matrix. The conditional $p(\mathbf{u}_h|\mathbf{y})$ is given as above and as such the multi-dimensional integral above provides an analytic Gaussian predictive distribution of the form,

$$p(\mathbf{y}_n|\mathbf{y}) = \mathcal{N}(\rho\Phi_n\mathbf{m}_{\mathbf{u}|\mathbf{y}}, \mathbf{K}_n + \sigma_\epsilon^2\mathbf{I}_n + \rho^2\Phi_n\mathbf{C}_{\mathbf{u}|\mathbf{y}}\Phi_n^\top). \quad (16)$$

It can be seen that the mean is the ρ -scaled interpolation of the updated FE model mean to the probe points in the domain, and the covariance is the sum of all the components contributing to the overall uncertainty in the prediction.

4.4 Obtaining the Full Marginal Distributions for Model Scoring

Denoting the full set of free random parameters as $\Omega = \{\theta, \sigma_f, l, \rho, \sigma_\epsilon\}$ recall that $p(\mathbf{y})$ is shorthand for $p(\mathbf{y}|\theta, \sigma_f, l, \rho, \sigma_\epsilon) \equiv p(\mathbf{y}|\Omega)$. A prior distribution can be defined over these random variables $\pi(\Omega)$ and the probability of the data given the choice of FE model and data generating model is obtained from,

$$p(\mathbf{y}|f, h) = \int p(\mathbf{y}|f, h, \Omega)\pi(\Omega)d\Omega. \quad (17)$$

Now given a specific forcing f , the two FE model representations, with resolutions h_1 and h_2 respectively, can be compared via,

$$\frac{p(\mathbf{y}|f, h_1)}{p(\mathbf{y}|f, h_2)} = \frac{\int p(\mathbf{y}|f, h_1, \Omega)\pi(\Omega)d\Omega}{\int p(\mathbf{y}|f, h_2, \Omega)\pi(\Omega)d\Omega}, \quad (18)$$

for example. These integrals are notoriously challenging to obtain numerical estimates of; Thermodynamic Monte Carlo Integration could be used however to implement this estimation ([Gelman and Meng, 1998](#)).

4.5 Multiple Sets of Observations

In the case where there are multiple sets of observations $\mathbf{y}_i$ (assumed to be at the coordinates $\mathbf{x}_y$, 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 posterior of $\mathbf{u}_h$ given all sets of observations is, $p(\mathbf{u}_h|\mathbf{Y}) = \mathcal{N}(\mathbf{m}_{\mathbf{u}|\mathbf{Y}}, \mathbf{C}_{\mathbf{u}|\mathbf{Y}})$, where

$$\begin{aligned} \mathbf{m}_{\mathbf{u}|\mathbf{Y}} &= \mathbf{C}_{\mathbf{u}|\mathbf{Y}} \left(\rho^2 \Phi^\top (\mathbf{K} + \sigma_\epsilon^2 \mathbf{I})^{-1} \sum_{i=1}^L (\mathbf{y}_i / \rho) + \mathbf{C}_{\mathbf{u}}^{-1} \mathbf{m}_{\mathbf{u}} \right) \\ \mathbf{C}_{\mathbf{u}|\mathbf{Y}} &= (\mathbf{C}_{\mathbf{u}}^{-1} + \rho^2 L \Phi^\top (\mathbf{K} + \sigma_\epsilon^2 \mathbf{I})^{-1} \Phi)^{-1}. \end{aligned}$$

It can be seen that when all the sets of observations $\mathbf{y}_i \sim \pi$ where π is some multivariate distribution with mean μ_π and if $\sigma_\epsilon^2 < \infty$, the posterior mean tends towards $\rho^{-1}\mu_\pi$ as L increases. Note that the number of repeated measurements does not impact the scaling of the number of Galerkin solves required, instead this is defined simply by the number of actual sensors $\mathcal{O}(N)$.

5 Application to the Mechanics of Thin Shell Structures

In this section we demonstrate the application of the proposed methodology to thin shell structures by means of a numerical example. 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.

5.1 Problem Description

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); Abueidda et al. (2017). 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).

The length of the beam is 0.8, its height, i.e. distance between the top and bottom plates, is 0.1 and its width is 0.1. The gyroid core is described by the algebraic function

$$\sin(\lambda x) \cos(\lambda y) - \sin(\lambda y) \cos(\lambda z) - \sin(\lambda z) \cos(\lambda x) = 0 \quad (19)$$

with $\lambda = 20\pi$. The core and the two plates are modelled as thin shells and have a thickness of $t = 0.003$. The beam is clamped at its left end, and at its right end the bottom plate is simply supported at a distance 0.025 away from the boundary. The Young's modulus and the Poisson's ratio are $E = 10^9$ and $\nu = 0.3$. The top plate is subjected to a uniform pressure $f(x) = 3 \cdot 10^6$ acting in the negative z direction. The uncertainty in the loading is modelled with a square exponential covariance function with a lengthscale $\sigma_f^2 = 10^{-4}$ and the covariance 10^{12} . The mathematical model considered in the following is linear so that the absolute value of the pressure and the associated probabilistic variables are not important. Therefore, they have been scaled so that the obtained displacements are comparable to the overall size of the beam.

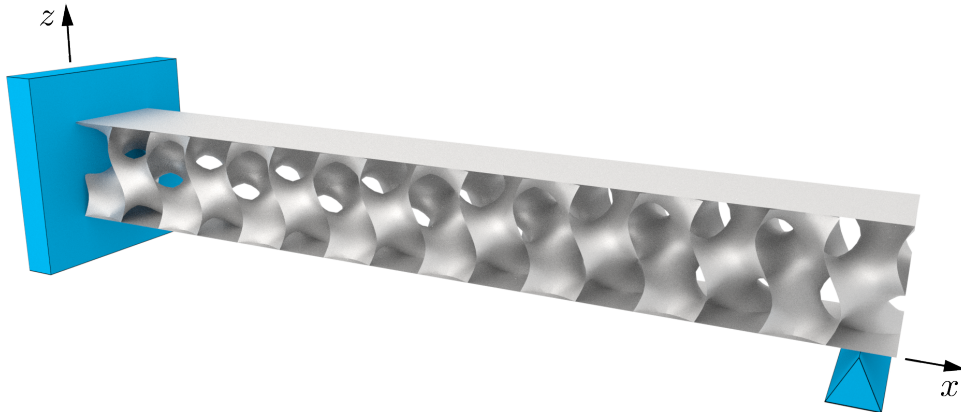


Figure 1: Schematic of the composite beam with a gyroid core and two face plates. The left end is clamped and right end is supported on a roller support. A uniform pressure $f(x)$ acting in the negative z direction is applied across the top plate.

5.2 Thin-Shell Model and FE Discretisation

The mechanics of thin shells can be conveniently analysed with a surface model using differential geometry (Ciarlet, 2005). According to the classical Kirchhoff-Love model, the mechanics of a shell surface depends on the change of its first and second fundamental forms during deflection with a displacement $u(x)$. The change in the first fundamental form yields the membrane strain tensor $\alpha(u(x))$ and the change in the second fundamental form yields the bending strain tensor $\beta(u(x))$. Both strain tensors have to be linearised to obtain a linear mathematical model. The bilinear form of the shell is additively composed of a membrane and a bending contribution,

$$\mathcal{B}(u(x), v(x)) = \mathcal{B}_m(u(x), v(x)) + \mathcal{B}_b(u(x), v(x)),$$

with,

$$\begin{aligned}\mathcal{B}_m(u(x), v(x)) &= \frac{Et}{1-\nu^2} \int \alpha(u(x)) : H(x) : \alpha(v(x)) j(x) \, d\Omega \\ \mathcal{B}_b(u(x), v(x)) &= \frac{Et^3}{12(1-\nu^2)} \int \beta(u) : H(x) : \beta(v(x)) j(x) \, d\Omega,\end{aligned}$$

where E is the Young's modulus and ν is the Poisson's ratio describing the elastic material, t is the thickness of the shell, and $H(x)$ is a fourth-order tensor and $j(x)$ is a scalar both depending only on the geometry of the non-deflected surface. Here $:$ denotes the double contraction of the two tensors. The linear form for the shell reads,

$$\langle f(x) + \epsilon_f(x), v(x) \rangle = \int \langle f(x) + \epsilon_f(x), v(x) \rangle \mu \, d\Omega.$$

It is worth emphasising that the integrals in the linear and bilinear forms are over the surface Ω , i.e. the two-manifold embedded in $\mathbb{R}^3$, which represents the solid shell structure with a small thickness.

As discussed in Section 2, for finite element discretisations, the functions $u(x)$ and $v(x)$ in the linear and bilinear forms are approximated as expansions of basis functions $\phi_i(x)$. The presence of second derivatives of $u(x)$ and $v(x)$ in the shell bilinear form requires that the approximants $u_h(x)$ and $v_h(x)$ are both square integrable. In the present implementation the basis functions are smooth quartic box splines that are defined on triangular meshes with certain symmetry and no boundaries. To be able to deal with boundaries and unstructured triangular control meshes we use the Loop subdivision surfaces which provide a generalisation of box splines; see Cirak et al. (2000); Cirak and Ortiz (2001); Cirak and Long (2011) for further details.

5.3 FE Model of the Composite Beam

The coarsest finite element control mesh, henceforth referred to as mesh 0, consists of 1088 triangular elements and 541 nodes. In the coarsest finite element mesh the nodes are placed on the gyroid surface given by (19). Finer meshes are obtained by successive refinement of the mesh using the Loop subdivision scheme (Zorin and Schröder, 2000). That is, the same basis functions $\phi_i(x)$ are used for generating the geometry and the FE discretisation, which is also referred to as isogeometric FE analysis (Hughes et al., 2005). In each refinement step each of the triangles are split into four triangles. Hence, the obtained refined control meshes have 4352, 17408 and 69632 elements, see Figure 2. These meshes are referred to as mesh 1, 2 and 3 respectively. All the control meshes describe the same smooth spline surface depicted in Figure 1, and the unknowns in the FE analysis are the displacements of the nodes of the control meshes. Furthermore, in FE the beam is composed of three distinct two-manifold surfaces representing the plates and the gyroid. The three two-manifolds are for mechanical purposes rigidly connected by enforcing the continuity of displacements and surface normals along common intersection points during the FE solution procedure. Boundary conditions are applied by constraining the displacements of relevant nodes to zero.

The deflected geometry of the composite beam after loading is depicted in Figure 3. It can be seen that the overall deflection of the entire plate is accompanied by a wavelike deflection of the relatively flexible top plate subjected to pressure loading. In the displacement plot in Figure 4 and subsequent plots, the

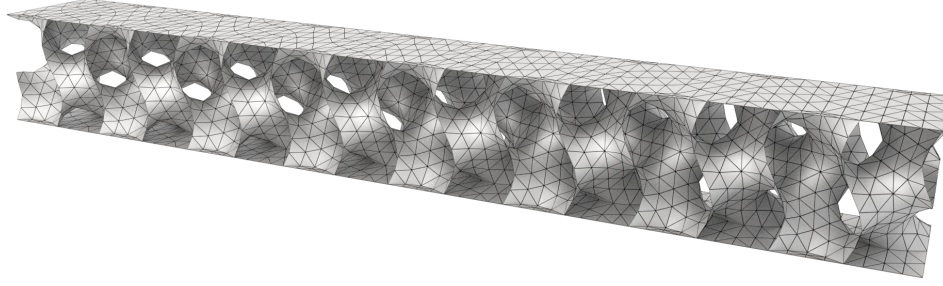


Figure 2: The finite element control mesh with 4352 triangular elements.

vertical displacement of the beam along the x -axis of the beam is always shown. The convergence of the mean of the displacements $\mathbf{u}_h$ and the measure of $\mathbf{u}_h$ with increasing mesh refinement is evident and the both quantities for the meshes with 17408 and 69632 elements are visually nearly indistinguishable.

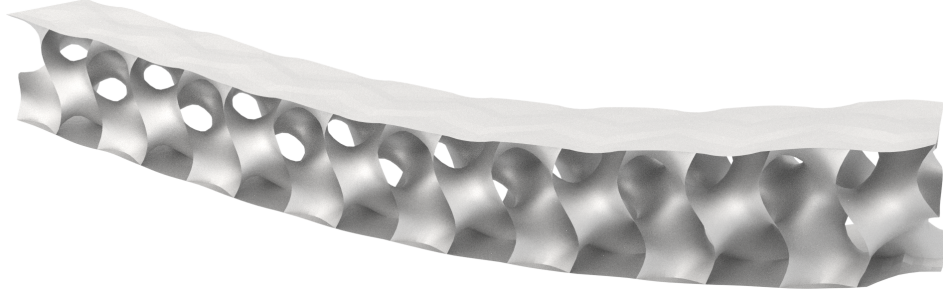


Figure 3: Deflected composite beam.

5.4 Modelling and Estimating the Mismatch Between Simulated Data and FE Model

In this example we use synthetic data which is assumed to be given at the points $\mathbf{x}_y$ by the formula,

$$\mathbf{y} = \boldsymbol{\eta} + \boldsymbol{\epsilon}, \quad (20)$$

where $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma_\epsilon^2 \mathbf{I})$ with $\sigma_\epsilon = 0.002$. The true data-generating deflection $\boldsymbol{\eta}$ is assumed to be given by the solution of the FE model, in the absence of noise. As described in Sec. 3, we assume that this data-generating process $\boldsymbol{\eta}$ is mismatched from $\mathbf{m}_u$ via,

$$\boldsymbol{\eta} = \rho \Phi \mathbf{m}_u + \boldsymbol{\delta},$$

where $\boldsymbol{\delta} \sim \mathcal{N}(\mathbf{0}, \sigma_\delta^2 \mathbf{K})$ is modelled by a mean-zero Gaussian measure, and $\rho = 0.7$. The covariance matrix $\mathbf{K}$ is a squared exponential covariance function with lengthscale $l = 0.4$ and scaling $\sigma_\delta = 0.07$ evaluated at $\mathbf{x}_y$. The points that we choose to observe the simulated data at are the sensor locations $\mathbf{x}_y$. There are in total 29 sensors, which are located along the x -axis of the beam. The sensors are uniformly distributed in between $\mathbf{x}_y \in [0.05, 0.075]$.

To generate the synthetic data a single realisation of the mismatch $\boldsymbol{\delta}$ is sampled, using the given parameters, and the data-generating process $\boldsymbol{\eta}$ is computed using the given value of ρ . Then, using $\mathbf{y} = \boldsymbol{\eta} + \boldsymbol{\epsilon}$, the data is generated at the coordinates $\mathbf{x}_y$. Figure 5 shows this data, generated using the mesh with 17408 elements (mesh 2) for the FE model. These quantities are superimposed on the measure of $\mathbf{u}_h$. Also shown in this figure is the conditional measure $p(\mathbf{u}_h | \mathbf{y})$, which represents the updated FE model once the data has been observed.

The parameters l and ρ , that are used to simulate the data $\mathbf{y}$, are assumed to be unknown. Estimates of the parameters l and ρ are obtained using simulated annealing (Brooks and Morgan, 1995) to maximise

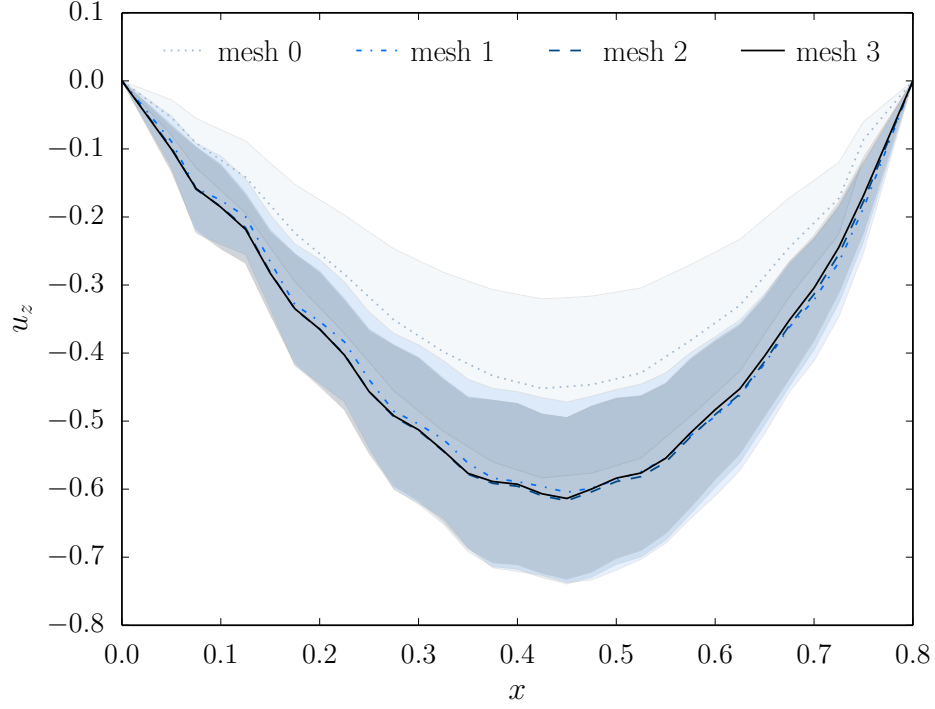


Figure 4: The mean and 95% confidence intervals of $\mathbf{u}_h$ for the four different meshes.

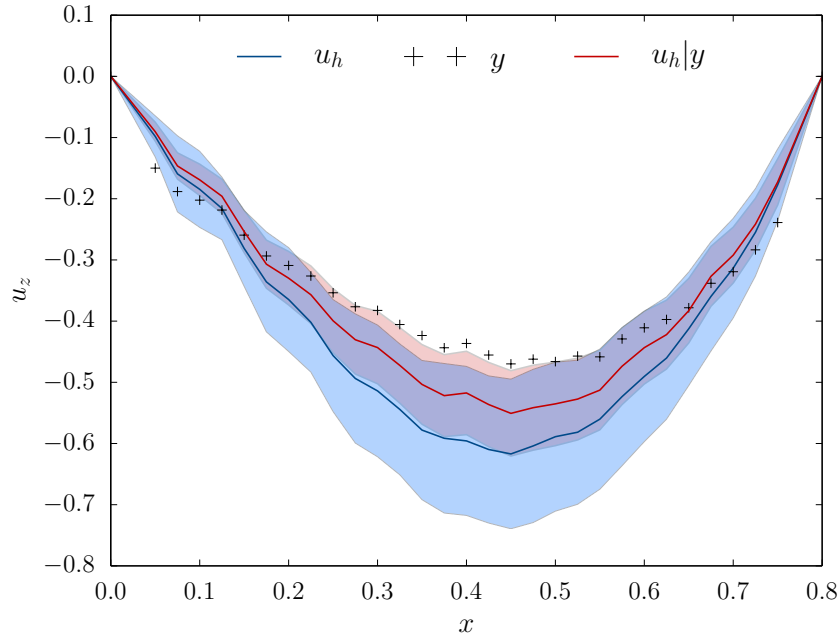


Figure 5: The measure of $\mathbf{u}_h$ (blue), in addition to the data (+) and the conditional measure $p(\mathbf{u}_h|\mathbf{y})$ (red).

the marginal likelihood in (6), with 50000 iterations and a cooling-rate of 3. The marginal densities of the Markov chains obtained by simulated annealing for each parameter are shown in Figure 6 respectively; the

final estimates of the two parameters, $\bar{l}$ and $\bar{\rho}$, are $\bar{l} \approx 0.4$ and $\bar{\rho} \approx 0.7$. The initial values of the parameters used as an input to the optimization routine are indicated as dashed lines in Figure 6.

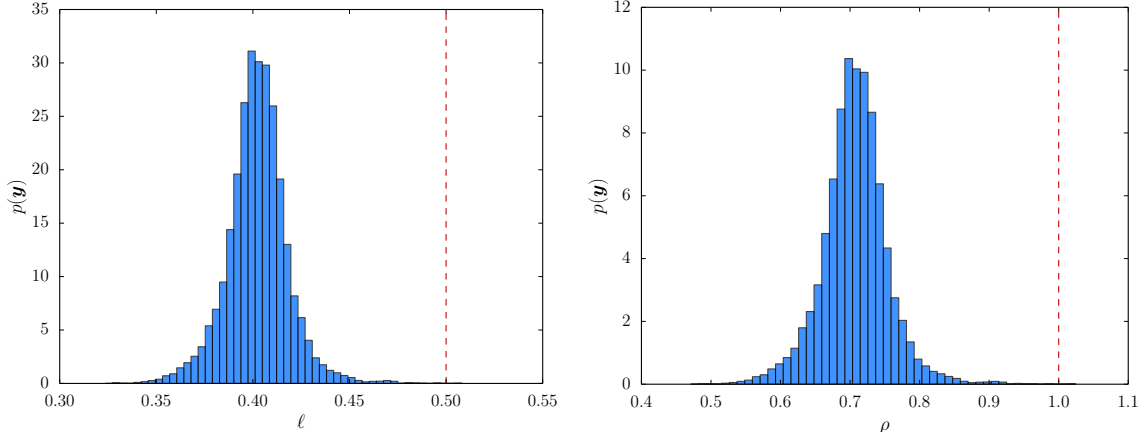


Figure 6: The marginal densities of the Markov chain for l and ρ (black), with the initial parameter values in red dashed lines.

5.5 FE Model Comparison

In this section we will compare the fit of different FE models with a given set of data; this is done using the Bayes factor described in Sec. 4.4. Three models are considered; the refinement of the meshes are the only thing that differentiates them. The considered meshes have 1088, 4352 and 17408 elements (meshes 0, 1 and 2). Data is generated, as done in the previous section, from a misspecification to the mean of each FE model measure $p(\mathbf{u}_h)$. These data will be denoted by $\mathbf{y}^{(0)}$, $\mathbf{y}^{(1)}$ and $\mathbf{y}^{(2)}$ respectively. Now one can compute the marginal likelihood in (6) for either set of data conditional on either FE model, e.g. $p(\mathbf{y}^{(i)}|f, h_j)$, for $i = 0, 1, 2$, and $j = 0, 1, 2$.

The parameters l and ρ that were assumed unknown and estimated in the previous section are now assumed known, so that the Bayes factor, described in Sec. 4.4, becomes simply $p(\mathbf{y}^{(i)}|f, h_i)/p(\mathbf{y}^{(i)}|f, h_j)$ for the i 'th set of data $\mathbf{y}^{(i)}$ with respect to the i 'th and j 'th FE model. The computed Bayes factors will now be used to assess the quality of the marginal likelihood as a metric to compare FE models with different mesh resolutions. The means (μ) and standard deviations (σ) of the log marginal likelihoods $\log p(\mathbf{y}^{(i)}|f, h_j)$, over 100 random data sets for each FE model, are given in the table below.

	$\mathbf{y}^{(0)}$	$\mathbf{y}^{(1)}$	$\mathbf{y}^{(2)}$
h_0	$\mu=95.5$; $\sigma=10.0$;	$\mu=55.2$; $\sigma=15.2$;	$\mu=55.3$; $\sigma=17.0$;
h_1	$\mu=77.6$; $\sigma=12.1$;	$\mu=99.0$; $\sigma=7.5$;	$\mu=96.7$; $\sigma=7.6$;
h_2	$\mu=72.2$; $\sigma=14.5$;	$\mu=95.1$; $\sigma=8.9$;	$\mu=97.0$; $\sigma=8.6$;

Note that the marginal likelihoods $p(y^{(i)}|f, h_i)$ are greater than the marginal likelihoods $p(y^{(i)}|f, h_j)$, for $j \neq i$, since the data is simulated from a different FE model than the model that the marginal likelihood is taken with respect to. Also apparent is the smaller differences between the marginal likelihoods using the data generated from the finer two models than the marginal likelihoods using the data generated from the coarsest model, likely because the FE models for these two meshes are very close together (see Figure 4).

6 Discussion and Future Research Directions

The statistical construction of finite element (FE) models, that provides a synthesis of a (possibly misspecified) mathematical model and measured data, has been presented in this paper. It is the first fully

statistical finite element method (FEM) that provides a data defined probability distribution capturing all sources of uncertainty in the twinning of FE models and sensor measurements. Further work will include addressing the uncertainty in the actual operator when θ is treated as random, and there has been much work employing multi-level Monte Carlo schemes that can be employed here; see [Teckentrup et al. \(2013\)](#) for more details.

Time evolving systems, and nonlinear PDEs fit into this framework naturally. For example, consider nonlinear PDEs of the form $\nabla \cdot (\varphi(u(x))\nabla u(x)) = f$ with Dirichlet boundary conditions. Denoting the nonlinear operator as $\mathcal{L}_{\varphi(u(x))} \equiv \nabla \cdot \varphi(u(x))\nabla \cdot$, then for the case where the right hand side is contaminated by random error and the weak form follows as $\langle \mathcal{L}_{\varphi(u(x))} u(x), v(x) \rangle = \langle f(x), v(x) \rangle + \langle \epsilon_f(x), v(x) \rangle$. The FEM representation then gives $\sum_{i,j \in I_h} u_i \langle \mathcal{L}_{\varphi(u_h(x))} \phi_i(x), \phi_j(x) \rangle = \langle f(x), \phi_j(x) \rangle + \langle \epsilon_f(x), \phi_j(x) \rangle$. The finite-dimensional conditional measure over the FE coefficients is then proportional to,

$$p(\mathbf{u}_h | \mathbf{A}, f, h) \propto \exp \left(-\frac{1}{2} \|\mathbf{A}_{\mathbf{u}_h} \mathbf{u}_h - \mathbf{b}\|_{\mathbf{G}^{-1}}^2 \right),$$

where the elements of the matrix $\mathbf{A}_{\mathbf{u}_h}$ are $\langle \mathcal{L}_{\varphi(u_h(x))} \phi_i(x), \phi_j(x) \rangle$. A local second order saddle-point approximation for the FE model distribution follows by defining a Gaussian over $\mathbf{u}_h$, whose mean $\mathbf{m}_{\mathbf{u}}$ solves for $\mathbf{A}_{\mathbf{u}_h} \mathbf{u}_h = \mathbf{b}$, and the precision matrix $\mathbf{C}_{\mathbf{u}}^{-1}$ has elements $\partial_i (\mathbf{A}_{\mathbf{u}_h} \mathbf{u}_h) \mathbf{G}^{-1} \partial_j (\mathbf{A}_{\mathbf{u}_h} \mathbf{u}_h)$ computed at $\mathbf{m}_{\mathbf{u}}$ with $\partial_i = \partial / \partial u_i$. The work presented in this paper is an important addition to the data-driven toolbox for FEM, which will find application in all of engineering and scientific applications where measurement rich systems are being studied and mathematical models are available.

References

- 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 & Design*, 122:255–267.
- Brooks, S. P. and Morgan, B. J. T. (1995). Optimization using simulated annealing. *The Statistician*, pages 241–257.
- Ciarlet, P. G. (2005). *An Introduction to Differential Geometry with Applications to Elasticity*. Springer.
- 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.
- 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.
- 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.
- 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.
- Gelman, A. and Meng, X.-L. (1998). Simulating normalizing constants: From importance sampling to bridge sampling to path sampling. *Statistical science*, pages 163–185.
- Ghanem, R. G. and Spanos, P. D. (2003). *Stochastic finite elements: a spectral approach*. Courier Corporation.
- Grafe, H. (1998). *Model updating of large structural dynamics models using measured response functions*. PhD thesis.
- Gregory, A., Lau, F. D.-H., Girolami, M., Butler, L. J., and Elshafie, M. Z. E. B. (2019). The synthesis of data from instrumented structures and physics-based models via gaussian processes. *Journal of Computational Physics*, 392:248–265.

- Hughes, T. J. R., Cottrell, J. A., and Bazilevs, Y. (2005). Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement. *Computer Methods in Applied Mechanics and Engineering*, 194:4135–4195.
- 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.
- Kennedy, M. C. and O’Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464.
- Strang, G. and Fix, G. J. (1973). *An analysis of the finite element method*, volume 212. Prentice-hall Englewood Cliffs, NJ.
- Stuart, A. M. (2010). Inverse problems: a bayesian perspective. *Acta numerica*, 19:451–559.
- Teckentrup, A. L., Scheichl, R., Giles, M. B., and Ullmann, E. (2013). Further analysis of multilevel monte carlo methods for elliptic pdes with random coefficients. *Numerische Mathematik*, 125(3):569–600.
- Zorin, D. and Schröder, P. (2000). Subdivision for modeling and animation. SIGGRAPH 2000 Course Notes.