P-1223

BAYESIAN OPTIMAL EXPERIMENTAL DESIGN FOR BIOMECHANICAL MODELS

Matthew Bilton[1], Thiranja P. Babarenda Gamage[2], Martyn Nash[1,2], Ruanui Nicholson[1], Oliver Maclaren[1]

[1]Department of Engineering Science, University of Auckland, New Zealand [2]Auckland Bioengineering Institute (ABI), The University of Auckland, New Zealand

THE UNIVERSITY OF AUCKLAND

Key Points

- We propose a flexible Bayesian Optimal Experimental Design (OED) workflow that can be applied to a wide range of biomechanical models; this workflow utilises Gaussian Process (GP) surrogate models and a Gaussian approximation to the posterior distribution.
- We use this workflow to show that the optimal angle at which to orientate a neo-Hookean cantilever beam to infer its stiffness by observing its deformation under gravity is 105°, not 90° as one might expect.
- We show that, in general, both the optimal design and the effect of the prior distribution on the optimal design are model-specific.

Introduction

All biomechanical models contain parameters, such as stiffness values, which must be estimated from experimental observations. However, it is not immediately obvious which experiment one should perform to gain the most information about a set of unknown parameters. The statistical field of Optimal Experimental Design (OED) provides a formal solution to this problem: perform the experiment which maximises an appropriately chosen 'experimental informativeness' metric. One popular OED metric is the *Expected Information Gain* (EIG) [1, 2]:

$$\mathrm{EIG}(d) = \mathbb{E}_{(\theta,y)\sim p(\theta,y|d)}\left[\ln p(\theta|y,d)\right] = \int p(\theta,y|d)\ln p(\theta|y,d)\,d\theta\,dy \tag{1}$$

Here, θ denotes the unknown parameters we wish to infer, d is a parameter that uniquely describes a particular experiment, y denotes the observations made after performing an experiment, and $p(\theta|y,d)$ is the posterior distribution. Within Bayesian statistics, the posterior distribution $p(\theta|y,d)$ quantifies one's beliefs about θ after having performed the experiment d and observing y; it is computed using Bayes' rule:

$$p(\theta|y,d) = \frac{p(y|\theta,d)\,p(\theta)}{p(y|d)} = \frac{p(y|\theta,d)\,p(\theta)}{\int p(y|\theta,d)\,p(\theta)\,d\theta} \tag{2}$$

Here, $p(y|\theta,d)$ is the *likelihood* and $p(\theta)$ is the *prior*, which describes our beliefs about θ before performing an experiment. Unfortunately, evaluating the EIG requires us to average over a large number of posterior distributions, which is very computationally expensive.

In this work, we propose a workflow to efficiently identify the EIG-maximising experiment for a wide range of biomechanical models. This workflow utilises *Gaussian Process* (GP) surrogate models in conjunction with a Gaussian approximation to the posterior distribution to compute the EIG efficiently.

Computation of Expected Information Gain

To help us efficiently compute the EIG, we shall make two simplifying assumptions:

1. Our observations y are formed by the outputs of a deterministic biomechanical model $f(\theta,d)$ being corrupted by zero-mean additive Gaussian noise with covariance Γ_y:

$$y = f(\theta,d) + \xi, \text{ where } \xi \sim \mathcal{N}(0,\Gamma_y) \tag{3}$$

Under this assumption, the likelihood takes the form of a Gaussian centred at $f(\theta,d)$ with covariance Γ_y:

$$y \sim \mathcal{N}\big(f(\theta,d),\Gamma_y\big) \implies p(y|\theta,d) \propto \exp\left[-\frac{1}{2}\big(y-f(\theta,d)\big)^T\Gamma_y^{-1}\big(y-f(\theta,d)\big)\right] \tag{4}$$

2. Our prior distribution is Gaussian with mean μ_θ and covariance Γ_θ:

$$p(\theta) \propto \exp\left[-\frac{1}{2}(\theta-\mu_\theta)^T\Gamma_\theta^{-1}(\theta-\mu_\theta)\right] \tag{5}$$

Under these assumptions, the posterior takes the form:

$$p(\theta|y,d) \propto \exp\left[-\frac{1}{2}\big(y-f(\theta,d)\big)^T\Gamma_y^{-1}\big(y-f(\theta,d)\big) - \frac{1}{2}(\theta-\mu_\theta)^T\Gamma_\theta^{-1}(\theta-\mu_\theta)\right] \tag{6}$$

Importantly, Equation (6) is **not** a Gaussian distribution with respect to θ if the biomechanical model $f(\theta,d)$ is non-linear. However, if our model $f(\theta,d)$ is linearised about the *maximum a posteriori* (MAP) point θ_{MAP}, like so:

$$f(\theta,d) \approx f(\theta_{\mathrm{MAP}},d) + \big(\partial_\theta f(\theta_{\mathrm{MAP}},d)\big)^T(\theta-\theta_{\mathrm{MAP}}) = \underbrace{\big(\partial_\theta f(\theta_{\mathrm{MAP}},d)\big)^T}_{=G(y,d)}\theta + \underbrace{f(\theta_{\mathrm{MAP}},d) - \big(\partial_\theta f(\theta_{\mathrm{MAP}},d)\big)^T\theta_{\mathrm{MAP}}}_{=b(y,d)} \tag{7}$$

where the MAP point is defined simply as the θ value which maximises Equation (6):

$$\theta_{\mathrm{MAP}} = \theta_{\mathrm{MAP}}(y,d) = \underset{\theta}{\mathrm{argmax}}\,p(\theta|y,d) = \underset{\theta}{\mathrm{argmax}}\,\exp\left[-\frac{1}{2}\big(y-f(\theta,d)\big)^T\Gamma_y^{-1}\big(y-f(\theta,d)\big) - \frac{1}{2}(\theta-\mu_\theta)^T\Gamma_\theta^{-1}(\theta-\mu_\theta)\right] \tag{8}$$

then the posterior in Equation 6 may be approximated as a Gaussian distribution [3, Section 3.2.2] with covariance:

$$\Gamma_{\theta|y,d}(y,d) = \left(G(y,d)^T\Gamma_y^{-1}G(y,d) + \Gamma_\theta^{-1}\right)^{-1} \tag{9}$$

and mean:

$$\mu_{\theta|y,d}(y,d) = \Gamma_{\theta|y,d}(y,d)\left(G(y,d)^T\Gamma_y^{-1}\big(y-b(y,d)\big) + \Gamma_\theta^{-1}\mu_\theta\right) \tag{10}$$

Under this Gaussian approximation, the EIG may be estimated using a simple Monte Carlo average:

$$\mathrm{EIG}(d) \approx \frac{1}{N}\sum_{i=1}^{N}\ln\mathcal{N}\big(\theta_i \,|\, \mu_{\theta|y,d}(y_i,d), \Gamma_{\theta|y,d}(y_i,d)\big), \text{ where } (\theta_i,y_i) \sim p(\theta,y|d) \tag{11}$$

To speed up the evaluation of Equation (11), a computationally cheaper *surrogate model* can be used in place of the full biomechanical model $f(\theta,d)$. One particularly flexible type of surrogate is a *Gaussian Process* (GP) regression model [4, Chapter 5]. In addition to being faster to evaluate, GP surrogates also provide us with *model gradients* which we require to linearise $f(\theta,d)$, even when the original model doesn't return gradient predictions [5, Section 9.4].

Optimal Experimental Design Workflow

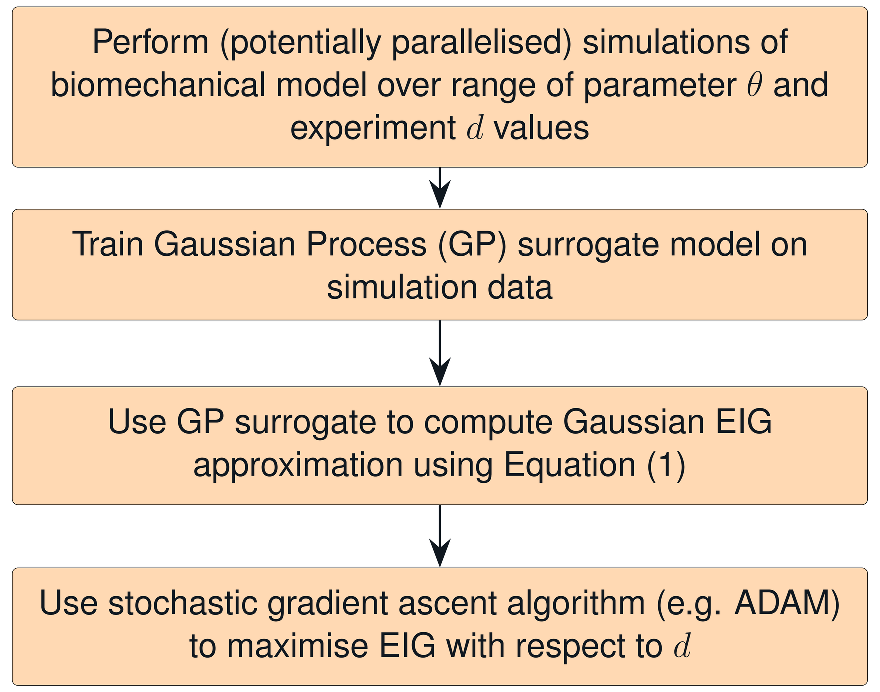

Fig. 1: Summary of proposed EIG-maximising OED workflow which utilises Gaussian process surrogates and a Gaussian approximation to the posterior distribution.

Cantilever Beam Toy Problem

To 'test' our proposed workflow, we chose to apply our OED methodology to the simple problem of identifying how one ought to orientate a soft cantilever beam to infer its stiffness by observing its deformation under gravity. More specifically, consider a nearly-incompressible neo-Hookean material with the following strain energy density Ψ:

$$\Psi(\boldsymbol{F}) = C_1\left(I_1(\boldsymbol{F}) - 3\right) + \kappa\left(J(\boldsymbol{F}) - 1\right)^2 \tag{12}$$

where C_1 is a stiffness parameter, κ is the bulk modulus which is approximately equal to 3000 kPa for soft tissues and silicon gels [6], F is the deformation gradient, $I_1(\boldsymbol{F})$ is the first invariant of the right Cauchy-Green deformation tensor, and J is the third invariant of the deformation gradient. Assuming that one can orientate the beam between an angle of $0°$ (i.e. the beam pointing straight down) and 180° (i.e. the beam point straight up) and then measure the displacement at the tip of the beam (see Figure 2), we wish to determine the best angle at which to orientate the beam to infer the stiffness parameter C_1.

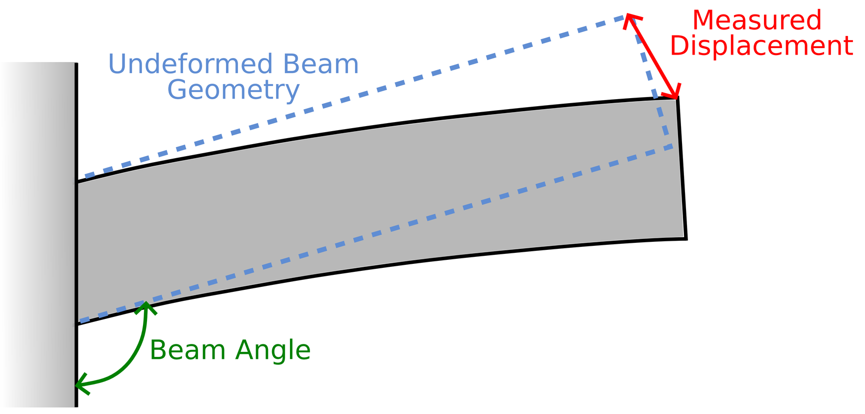

Fig. 2: Schematic illustration of the cantilever problem.

To generate training data for our GP surrogate, the Dolfinx package developed by the FEniCS Project [7] was used to simulate a neo-Hookean beam for 100 different C_1 and beam angle combinations. With our surrogate trained, the EIG could be estimated using the Gaussian approximation given by Equation (11), as shown in Figure 4; this landscape was compared to directly computing Equation (1) using numerical integration with the GP surrogate. Surprisingly, the optimal beam angle to infer C_1 is **not** 90° (i.e. holding the beam horizontally), but is instead around 105°.

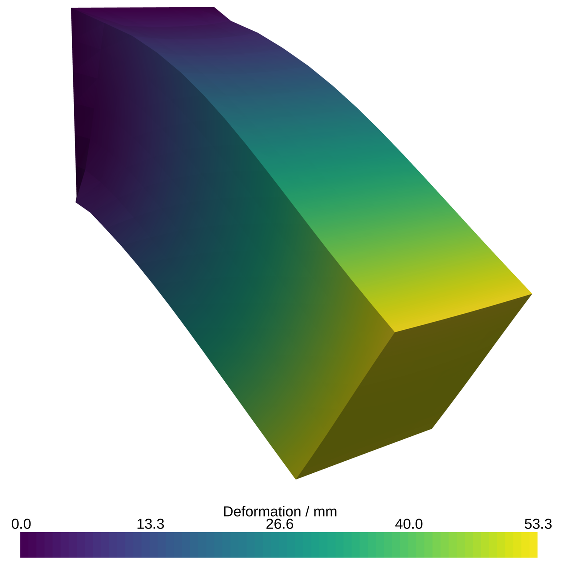

Fig. 3: Dolfinx model of nearly incompressible neo-Hookean beam with stiffness C_1 = 2 kPa, bulk modulus κ = 3000 kPa, and density $\rho = 1.02\times10^{-3}$ g·cm^{-3} deforming under gravity.

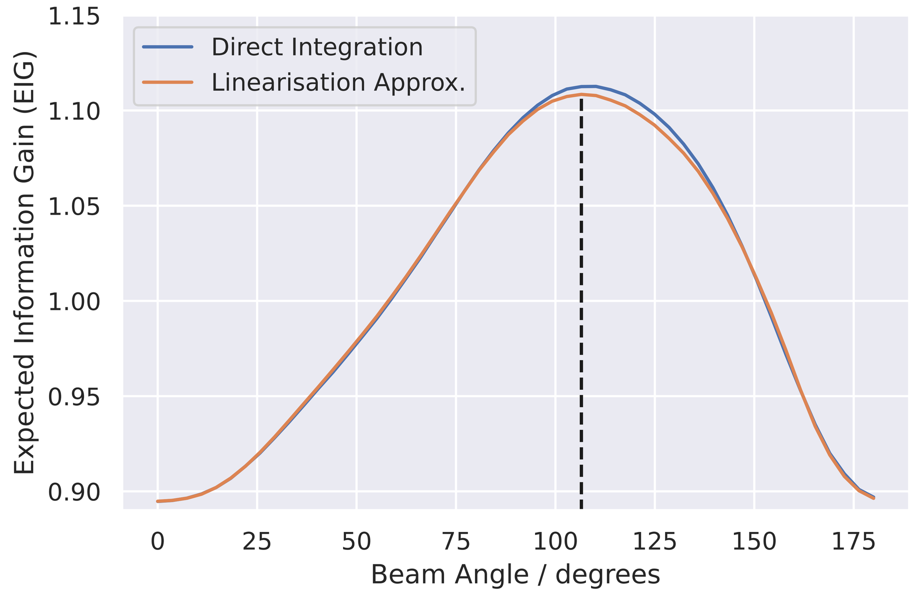

Fig. 4: Expected Information Gain (EIG) as a function of beam angle, plotted for μ_θ = 4.4 kPa, Γ_θ = 0.25 kPa^2, Γ_y = 16.9 mm^2. The optimal design is approximately $105°$ in this case.

Effect of Biomechanical Model & Prior on Optimal Design

In the cantilever toy problem, we were forced to choose a particular biomechanics model $f(\theta,d)$ and prior $p(\theta)$. But how do these decisions affect the optimal beam angle? To investigate this, an additional GP surrogate was created by simulating a *linear elastic* cantilever beam. Although the assumptions of linear elasticity are not satisfied here, our use of linear elasticity here should be viewed as an 'extreme case' of using an overly-simple mathematical model to describe a complex physical system. The EIG landscapes for both the neo-Hookean beam and the linear elastic beam were then computed for different μ_θ and Γ_θ values. Interestingly, the optimal angle for the linear elastic beam is $90°$ instead of 105°. Moreover, changing μ_θ does not affect the optimal angle of the linear elastic beam, but does for the neo-Hookean beam. Changing Γ_θ does not affect the optimal design of both cantilever beams.

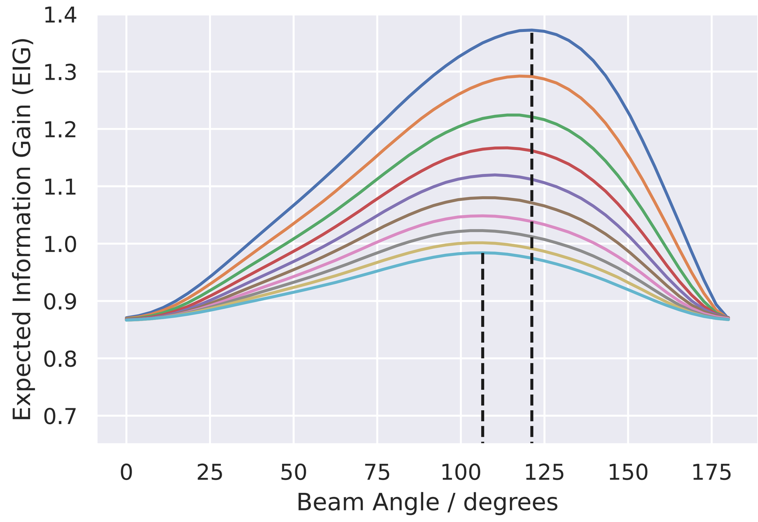

Fig. 5: Effect of changing prior mean μ_θ on EIG landscape for *neo-Hookean* cantilever beam problem for fixed Γ_θ = 0.25 kPa^2 and Γ_y = 16.9 mm^2.

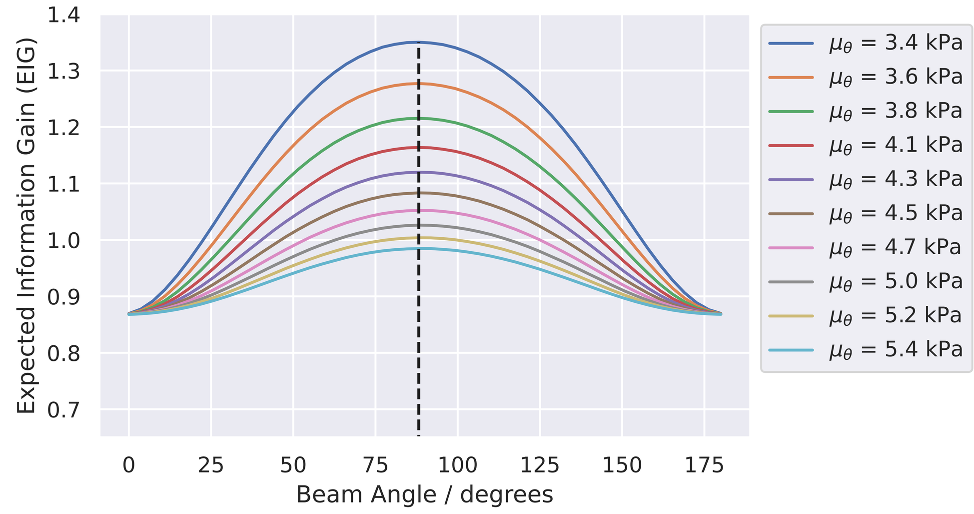

Fig. 6: Effect of changing prior mean μ_θ on EIG landscape for *linear elastic* cantilever beam problem for fixed Γ_θ = 0.25 kPa^2 and Γ_y = 16.9 mm^2.

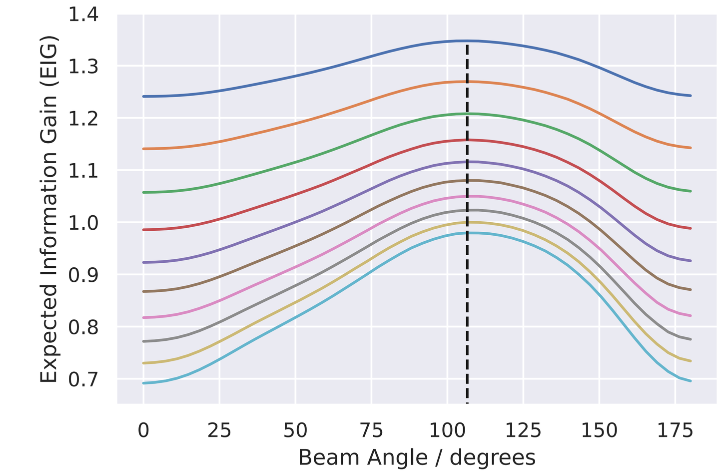

Fig. 7: Effect of changing prior covariance Γ_θ on EIG landscape for *neo-Hookean* cantilever beam problem for fixed μ_θ = 4.4 kPa^2 and Γ_y = 16.9 mm^2.

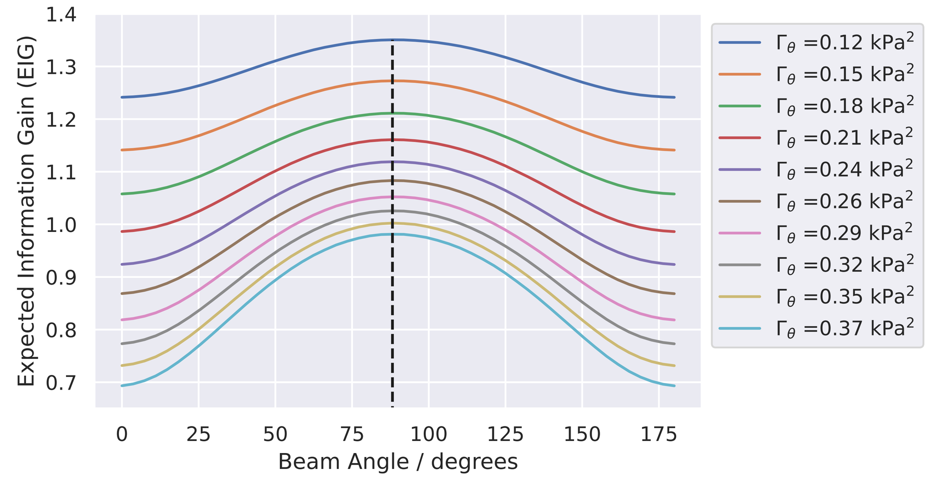

Fig. 8: Effect of changing prior covariance Γ_θ on EIG landscape for *linear elastic* cantilever beam problem for fixed μ_θ = 4.4 kPa^2 and Γ_y = 16.9 mm^2.

Conclusions

- An OED methodology to efficiently compute the Expected Information Gain (EIG) using GP surrogates and a posterior Gaussian approximation was described and applied to a cantilever beam toy problem.
- Surprisingly, the optimal angle at which to orientate a neo-Hookean cantilever beam to infer its stiffness is 105°, rather than 90°.
- Optimal designs are *model-specific*: whereas the optimal design for a neo-Hookean beam is 105° degrees, the optimal design for a linear elastic beam is 90° degrees.
- The effect of the prior on the optimal design is *also model specific*: the prior mean does not affect the optimal design of a linear elastic beam, but it does affect the optimal design of a neo-Hookean beam.

Code Availability & Acknowledgements

All of the code used to produce the results presented in this poster, in addition to the all other results presented in the main author's masters thesis, may be found at the following GitHub-hosted Git repository: `https://github.com/MABilton/bayesian_oed_bilton_masters`

Matthew Bilton would like to thank the University of Auckland for providing the Masters Research Scholarship that funded this work.

References

[1] Adam Foster et al. "Variational Bayesian optimal experimental design". In: *Advances in Neural Information Processing Systems* 32 (2019).

[2] EG Ryan et al. "Fully Bayesian optimal experimental design: A review". In: *International Statistical Review* 84 (2016), pp. 128–154.

[3] Albert Tarantola. *Inverse problem theory and methods for model parameter estimation.* SIAM, 2005.

[4] Robert B Gramacy. *Surrogates: Gaussian process modeling, design, and optimization for the applied sciences.* Chapman and Hall/CRC, 2020.

[5] Christopher KI Williams and Carl Edward Rasmussen. *Gaussian processes for machine learning.* Vol. 2. 3. MIT press Cambridge, MA, 2006.

[6] Kevin M Moerman et al. "Digital image correlation and finite element modelling as a method to determine mechanical properties of human soft tissue in vivo". In: *Journal of biomechanics* 42.8 (2009), pp. 1150–1153.

[7] Matthew W Scroggs et al. "Construction of arbitrary order finite element degree-of-freedom maps on polygonal and polyhedral cell meshes". In: *ACM Transactions on Mathematical Software (TOMS)* 48.2 (2022), pp. 1–23.