Utilising Surrogate Models to Approximate Cardiac Potentials when Solving Inverse Problems via Bayesian Techniques

Abbish Kamalakkannan[1], Peter Johnston,[1], Barbara Johnston,[1]

[1] Griffith University, Brisbane, Australia

Abstract

Solving inverse problems is computationally expensive, if not infeasible, under specific scenarios. For example, many forward solutions are required when solving inverse problems using Bayesian techniques. In this work, a novel inference protocol is established, that can be used to infer the cardiac bidomain conductivities and the cardiac fibre rotation angle (bidomain parameters). This protocol uses a surrogate model, developed using generalised polynomial chaos techniques, to approximate cardiac potentials on a multi-electrode array. The resulting surrogate model is used in conjunction with Bayesian inference techniques to infer the bidomain parameters. A lower-order surrogate model (order three) can effectively characterise the influence of the extracellular conductivities and fibre rotation on the cardiac potentials; however, it is recommended that a higher-order surrogate model expansion of order seven be used to adequately characterise the influence of the intracellular conductivities as well. This seventh order surrogate model was successfully used to infer the extracellular conductivities and fibre rotation angle from a single set of synthetically generated noisy experimental potentials, while the intracellular conductivities were unable to be retrieved accurately under this scenario.

1. Introduction

Electrophysiological simulations of cardiac tissue can be a valuable tool to help facilitate a more thorough understanding of illnesses of the heart. Such simulations can be achieved by using the bidomain equations, which govern the electrical potential in the intracellular i and extracellular e spaces of cardiac tissue, under the assumption that they are spatially coincident continua.

Heart tissue, unique in structure compared to other muscle tissue [1], is arranged in laminar sheets of fibres and stacked such that the orientation of each sheet is slightly offset to the sheet below to form the walls of the heart, thus introducing the notion of fibre rotation. The electrical conduction in the three orthogonal directions - along and across the direction of the fibres (longitudinal l and transverse t directions, respectively) and normal n to the sheets of the fibres - is different [2]. Given there are two interpenetrating domains, a total of six cardiac conductivity values are required to describe the distribution of potential through cardiac tissue [3]: the intracellular conductivities, g_{il}, g_{it}, g_{in}, abbreviated as g_i's, and the extracellular conductivities, g_{el}, g_{et} and g_{en}, abbreviated as g_e's. These six cardiac conductivity values and the fibre rotation angle α (the bidomain parameters) are critical for accurate electrophysiological simulations of cardiac tissue through the use of the bidomain model.

Difficulties associated with obtaining and utilising experimental measurements [3] have made it challenging to accurately obtain values for the bidomain parameters. This study develops a novel protocol, based on a Bayesian framework and polynomial chaos techniques, which allows for the creation of a surrogate model that approximates solutions to the bidomain model. This surrogate model is used to retrieve the bidomain parameters from experimental potentials measured on a multi-electrode array inserted into cardiac tissue. This protocol is validated through the use of simulated experimental measurements.

2. Methods

2.1. Governing Equations

In this model, we assume that the ventricular muscle is represented by a slab of cardiac tissue extending $2\,\mathrm{cm}$ in the x and y directions and $1\,\mathrm{cm}$ in the z direction, which is in contact with a pool of blood whose thickness extends to infinity. It is assumed that the l, t and n directions are aligned along the x, y and z axes, respectively, on the epicardium [3].

The electric potentials, within the cardiac tissue, are governed by the steady-state bidomain equations [2],

$$\boldsymbol{\nabla}\cdot\boldsymbol{M}_i\nabla\phi_i = \frac{\beta}{R}(\phi_i-\phi_e),\ \boldsymbol{\nabla}\cdot\boldsymbol{M}_e\nabla\phi_e = -\frac{\beta}{R}(\phi_i-\phi_e)-I_c\,,$$

$$(1)$$

where ϕ_q is the potential ($q = i, e$), R is the resistance of the cell membrane, β is the ratio of surface cell area to cell
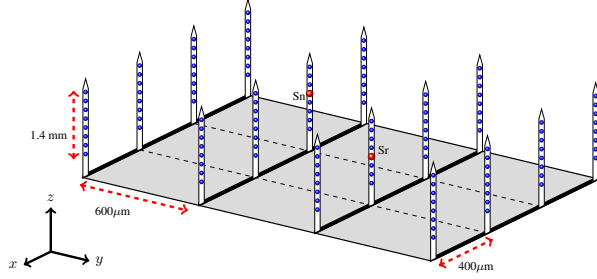
Figure 1. Schematic of the NeuroNexus128-electrode array. The current source and sink electrodes, labelled Sr and Sn, respectively, are coloured red.

volume, and I_c is the sub-threshold current applied in the extracellular space to produce a potential. M_q is a conductivity tensor that accounts for the anisotropy of cardiac tissue and contains the bidomain parameters including the fibre rotation angle α. Laplace's equation governs potential in the blood. A description of the solution technique is given in a previous paper by Johnston and Johnston [3].

The following parameters were used to solve the model: $\beta = 2000\text{cm}^{-1}$, $R = 9100\ \Omega\text{cm}^2$, $I_c = 1\ \mu\text{Acm}^{-3}$, with the conductivity of the blood assumed to be 6.7 mS/cm. These parameters are consistent with values used for electrophysiological simulations from the literature [3, 4]. The nominal values used for the bidomain parameters are $g_{il} = 2.40$, $g_{it} = 0.24$, $g_{in} = 0.10$, $g_{el} = 2.40$, $g_{et} = 1.60$, $g_{en} = 1.00$, $\alpha = 120°$ where the conductivities are in mS/cm [3].

2.2. Electrode Array and Generating Synthetic Potentials

A NeuroNexus Matrix Array (NeuroNexus, Ann Arbor, MI, USA) consisting of 128 electrodes (Figure 1) is used as the basis for this study. Two electrodes are used for the current application, while the remainder for measuring the resulting extracellular potentials.

Noise samples from a Gaussian distribution with a mean of zero and a specified standard deviation are added to the nominal synthetic potentials calculated on each measuring electrode via the bidomain model. The resulting set of potentials is labelled a measurement set and is equivalent to a single set of experimental measurements, given a pair of current injection electrodes.

2.3. Inference Protocol

Bayes' theorem is employed to infer the probability distributions of the seven bidomain parameters $\boldsymbol{\theta}$ (posterior), given data D, the forward model and prior knowledge about $\boldsymbol{\theta}$ [5]. Markov Chain Monte Carlo algorithms

(NUTS and Slice Samplers) are used to approximate the posterior distributions through the PYMC3 package [6].

2.3.1. The Likelihood Function

The likelihood function is constructed by first assuming the measurement error is given by [7],

$$D_w = f_w(\boldsymbol{\theta}) + \epsilon_w, \tag{2}$$

where D_w is the experimental data at the wth measuring electrode, $f_w(\boldsymbol{\theta})$ is the forward model potential at the wth measuring electrode, and ϵ_w is a normally distributed random variable that accounts for measurement noise introduced to the wth measuring electrode. The joint likelihood of the measurement set is the product of the individual observation likelihoods [8], that is,

$$P(D|\boldsymbol{\theta}, \sigma^2) = \prod_{w=1}^{W} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(D_w - f_w(\boldsymbol{\theta}))^2}{2\sigma^2}\right) \tag{3}$$

where W is the number of measuring electrodes, and σ^2 quantifies the experimental uncertainties of all forms. No prior experimental information exists regarding σ^2, and so it will be treated as a nuisance parameter and inferred during parameter estimation.

2.4. Introduction of a Surrogate Model

Given the computational costs of solving the bidomain model, it is often infeasible to infer the bidomain parameters through the Bayesian framework. However, a generalised polynomial chaos (gPC) expansion [9] can approximate the cardiac potentials at each measuring electrode, given $\boldsymbol{\theta}$. Simulations of the forward model can then be replaced by evaluations of the gPC expansion, which is computationally inexpensive.

The following gPC expansion is constructed using the forward model realisations obtained using N_p samples from the parameter space [9],

$$\hat{f}_{N,w}(\boldsymbol{\theta}) = \sum_{\eta=0}^{N_p-1} c_\eta \psi_\eta(\boldsymbol{Z}), \tag{4}$$

where N is the polynomial order, $\hat{f}_{N,w}$ is the approximation of the cardiac potential at the wth measuring electrode given $\boldsymbol{\theta}$ and N_p is the truncation term dependent on N and the number of model parameters. The polynomial expansion is constructed using Legendre polynomials (ψ_η) since the samples are uniformly distributed. The polynomial coefficients (c_η) are evaluated using regression techniques [9].

In this work, the gPC expansion is constructed to approximate the potentials using bidomain parameters in the

domain $[0.5M_0, 1.5M_0]$, where M_0 is the nominal value of the bidomain parameter. It is assumed that the prior distribution of each bidomain parameter in $\boldsymbol{\theta}$ is uniformly distributed from $[0.5M_0, 1.5M_0]$. The prior for σ is assumed to be given by a half-normal distribution, given that σ is continuous and strictly positive.

3. Results and Discussion

3.1. Identifying the Errors of the Surrogate Model

In this study, we attempt to identify any errors in the retrieved bidomain parameters that arise due to the introduction of the surrogate model. This can be achieved by inferring $\boldsymbol{\theta}$, using the surrogate model, from a nominal measurement set of potentials with no added noise. Note that nominal potentials are generated using the bidomain model at the location of the measuring electrodes and using the source-sink electrode given in Figure 1.

Firstly, an expansion of order three, with $N_p = 130$, is used to develop the surrogate model (low-fidelity model [9]). This surrogate model has an L_2 relative error norm of 0.840% when compared against 801 truth datasets. In addition, an expansion of order seven is also used to develop the surrogate model (mid-fidelity model), with $N_p = 3442$ [9]. The mid-fidelity surrogate model has an L_2 relative error norm of 0.020% when compared against 801 truth datasets. A sample takes approximately 0.4 seconds to be computed; however, once the surrogate has been created, the computational cost of evaluating the mid-fidelity model is comparable to that of the low-fidelity model.

When combined with the inference protocol, the low-fidelity model can retrieve the g_e's and α accurately. The means of the posterior distributions of these parameters have relative errors of approximately 0.3%. The 99% highest density interval (HDI), that is, the interval in which the retrieved parameter falls with probability 0.99 and which describes the uncertainty in the retrieved value, has bounds that are approximately $\pm 0.2\%$ of the mean value for the g_e's and $\pm 1\%$ of the mean value for α.

The relative errors of the mean values of the retrieved g_i posterior distributions were 1.37%, 3.45% and 5.32% for g_{il}, g_{it} and g_{in}, respectively, with 99% HDI bounds of $\pm 0.2\%$, $\pm 7.7\%$ and $\pm 5.6\%$, respectively.

Under the mid-fidelity model, the means of the retrieved g_i posterior distributions have a relative error of approximately $3.1 \times 10^{-3}\%$, with bounds of the 99% HDI being $\pm 0.03\%$ of the mean. The g_e distributions have a relative error of approximately $2.0 \times 10^{-3}\%$, and the 99% HDI bounds are $\pm 1.2 \times 10^{-3}\%$ of the mean value. Finally, the mean for the distribution of α has a relative error of $9.3 \times 10^{-3}\%$ and has 99% HDI bounds that are

$\pm 8.1 \times 10^{-3}\%$ of the mean value.

These results show that the low-fidelity forward surrogate model can accurately characterise g_e's and α, but is unable to adequately characterise the influence of g_i's, even given the low L_2 error of the approximation. Consequently, the g_i's were retrieved with higher errors when using the low-fidelity model. The mid-fidelity forward surrogate model can accurately characterise the influence of all the bidomain parameters on the cardiac potentials, and thus the bidomain parameters are able to be retrieved accurately. Given the reduced computational costs of developing the low-fidelity surrogate, it is recommended that the low-fidelity surrogate be used to refine the inference algorithm parameters, and the mid-fidelity surrogate model be used for parameter inference.

3.2. Introducing Simulated Experimental Noise

Synthetic experimental potentials are generated by introducing noise to a set of simulated nominal potentials. Noise samples are obtained from a normal distribution of mean 0 volts and a standard deviation of 1.5×10^{-5} volts (approximately 5% of the root mean square of the nominal potentials). A mid-fidelity surrogate model is used to infer $\boldsymbol{\theta}$ given the synthetically generated measurement set. Figure 2 presents the inference results.

Figure 2 suggests that convergence has been achieved for the g_e's and α. The relative error of the mean of the posterior distributions for the g_e's is approximately 1%, while the bounds of the 99% HDI are approximately $\pm 3.7\%$ of the mean value. Similarly, the mean of the distribution of α has a relative error of 5%, while the bounds of the 99% HDI are approximately $\pm 25\%$ of the mean. The g_i's are not adequately inferred, and the 99% HDI almost spans the domain of the gPC model. However, it should be noted that this protocol provides more desirable results than current frequentist approaches that are used to retrieve the bidomain parameters given a single set of synthetic measurements of similar noise levels [3].

Validation of the inference (posterior predictive check) can be achieved by generating replicated datasets, obtained using the forward model and samples of the bidomain parameters drawn from the inference, and comparing it with the observed data (the measurement set). The results of posterior predictive checks suggest that the inferred parameters are able to replicate the observed data very well, despite the high levels of uncertainty associated with the g_i's and the errors associated with the g_e's and α.

4. Conclusion

This study developed a novel protocol to retrieve the bidomain parameters through the use of a surrogate model
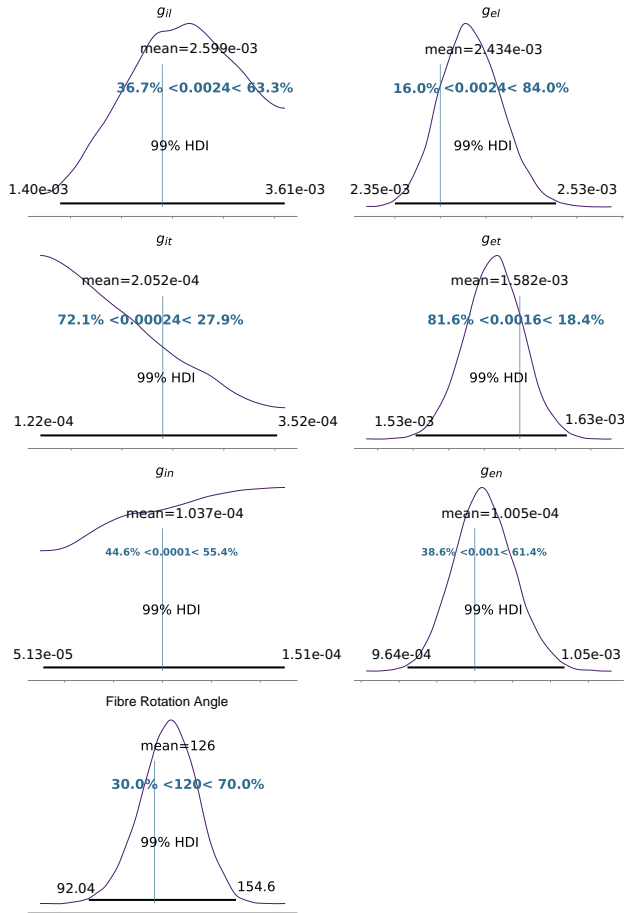
Figure 2. Posterior distributions, mean and the 99% HDI of the retrieved bidomain parameters. The vertical line indicates the nominal value of the parameter. The inequalities indicate the probability of the retrieved value either being greater or lesser than the nominal value. Values for the conductivities are in S/cm.

and Bayesian inference techniques. A lower-order surrogate model of order three can be used to accurately characterise the influence of the extracellular conductivities and fibre rotation on the solution; however, a higher-order expansion of order seven is required to accurately characterise the influence of the intracellular conductivities as well.

Using this order seven protocol to infer the bidomain parameters from a simulated noisy dataset suggests that extracellular conductivities and fibre rotation angle can be inferred to an acceptable degree of accuracy; however, the intracellular conductivities were unable to be inferred. We hypothesise that the use of multiple datasets, obtained through different source-sink electrode configurations, may allow for the accurate retrieval of intracellular conductivities, as well as allow for extracellular conductivities and fibre rotation angle to be further refined. Future studies will further develop this protocol to include data from multiple datasets and tailor the likelihood such that the presence of outliers does not unduly influence the results of the inference.

Acknowledgments

References

[1] Legrice IJ, Hunter PJ, Smaill BH. Laminar structure of the heart: A mathematical model. American Journal of Physiology Heart and Circulatory Physiology 1997; 272(5):H2466–H2476.

[2] Tung L. A bi-domain model for describing ischemic myocardial D-C potentials. Thesis, Massachusetts Institute of Technology, 1978.

[3] Johnston BM, Johnston PR. Approaches for determining cardiac bidomain conductivity values: Progress and challenges. Medical Biological Engineering Computing December 2020;58(12):2919–2935. ISSN 1741-0444.

[4] Plonsey R, Barr R. The four-electrode resistivity technique as applied to cardiac muscle. IEEE Transactions on Biomedical Engineering July 1982;29(7):541–546. ISSN 0018-9294.

[5] Gelman A, Carlin JB, Stern HS, Rubin DB. Bayesian data analysis. Chapman and Hall/CRC, 1995.

[6] Salvatier J, Wiecki TV, Fonnesbeck C. Probabilistic programming in python using PyMC3. PeerJ Computer Science 2016;2:e55.

[7] Kamalakkannan A, Johnston PR, Johnston BM. A modified approach to determine the six cardiac bidomain conductivities. Computers in Biology and Medicine 2021;.

[8] Bolstad WM, Curran JM. Introduction to bayesian statistics. John Wiley & Sons, August 2016. ISBN 978-1-118-59315-8.

[9] Xiu D. Fast numerical methods for stochastic computations: A review. Communications in Computational Physics 2009;31.

Address for correspondence:

Abbish Kamalakkannan, School of Environment and Science, Queensland Micro- and Nanotechnology Centre, Griffith University, Nathan, Queensland, Australia, a.kamalakkannan@griffith.edu.au