Chapter 9

Studying the effects of physiological conditions on model calibration

The end goal of model calibration in mathematical models of cardiac function is to provide real world information on cardiac health. By linking the underlying behaviour of the heart with observable quantities, tuning these models to match a given dataset can provide insight into the system at a level that is naked to the human eye. While fundamentally informative, the utility of our parameter estimation depends on the accuracy of the mathematical model: if the model is imperfect, we may only ever obtain a representation in the paramters that is imperfect. Such imperfections—which can be the result of measurement error or suboptimal modelling— are an inevitability of the translation of mathematical models to the real world. Acknowledging these imperfections necessitates the use of statitstical methodology that can correctly account for the discrepancy between the model and the data. This chapter will consider this task in the context of parameter estimation for data measured from cardiac magnetic resonance (CMR) imaging scans. In addition, the effect that operator errors during LV geometry segmentation have on the parameter estimation task will be explored. The LV operator study is joint with David Dalton and Arash Rabbani, who acted as the other two operators.

9.1 Model discrepancy background

9.1.1 Motivation with a set of measured data

Let us begin with an example that shows the typical results of parameter inference for measured LV data. Directly from CMR scans we can measure a set of 24 circumferential strains and the end diastolic volume of the cavity (LVV), as well as the LV geometry G, which has vector representation $\mathbf{g}$. The parameterization of the H-O model used for this study is the one proposed by the SA in Chapter 8. Adopting the likelihood

function used in Section 8.2.5, predictions of the LVV and strains are provided by an emulator $\boldsymbol{\eta}(\boldsymbol{\theta}, \mathbf{h})$, where $\mathbf{h}$ is the 5 PC representation of $\mathbf{g}$. The material parameters a, b, a_f and b_f can be sampled from the posterior distribution under an assumption of iid zero-mean Gaussian noise on the circumferential strains and a noise standard deviation on the volume held fixed at 5 ml, which is consistent with the results of an empirical study that is found in Appendix E.1. The parameter sampling is performed using NUTS, before thinning to get the final sample to be used for the analysis. In Figure 9.1, the posterior distributions of $\boldsymbol{\eta}(\boldsymbol{\theta}, \mathbf{h})$, obtained by predicting with the emulator at the thinned material parameter samples, are plotted as solid black density curves for each model output. The vertical dashed line, showing the measured value of that output, allows us to identify any discrepancy between the posteriors and the observations. For comparative purposes, the inference task is repeated with a likelihood function that contains only the LVV, with the results displayed in the same plots using dashed density curves. Recall that strain measurements are negative, so a density lying to the right of the dashed vertical line shows an underestimation of the strain. For the volume (shown in the top left plot), the opposite is true: density to the left of the vertical dashed line indicates an underestimation of the LVV.

The solid density lines in Figure 9.1 (which result from inference conditional on volume and strains) present a familiar theme, and one that was touched on in Section 7.4.4: in order to match the measured strains, we must underestimate the volume. The word "match" is used rather loosely here because, as we can see in the plot, many of the densities do not match the observed values. The level of noise in the measured data is so large that the measurements span the majority of the space of possible strain values (in some cases between -0.05 and -0.25). In the presence of such high levels of noise, the best we can do is match the mean strain and this is always in disagreement with the LVV measurement, which can only be matched at parameter configurations that correspond to predicted circumferential strains that are of greater (in absolute value) than 0.2.

In Figure 9.2, the measured circumferential strains for 8 different test cases are compared with simulated circumferential strains obtained at material parameter values that allow us to match the LVV. Comparing the observations and simulations in each test case presents two consistent findings. The first of these is the high degree of noise in the measured data. In the majority of test cases, the IQR of the measured data is significantly larger than that of the simulated data. The second finding is that the simulated circumferential strains correspond to a greater deformation than the measured circumferential strains. This pattern suggests that some combination of the following is occuring: the measurement noise is not iid zero-mean Gaussian (related to the measurement process) or there is a systematic mismatch between the simulator and underlying physical process (related to the model building process). Each of these problems can manifest in a similar
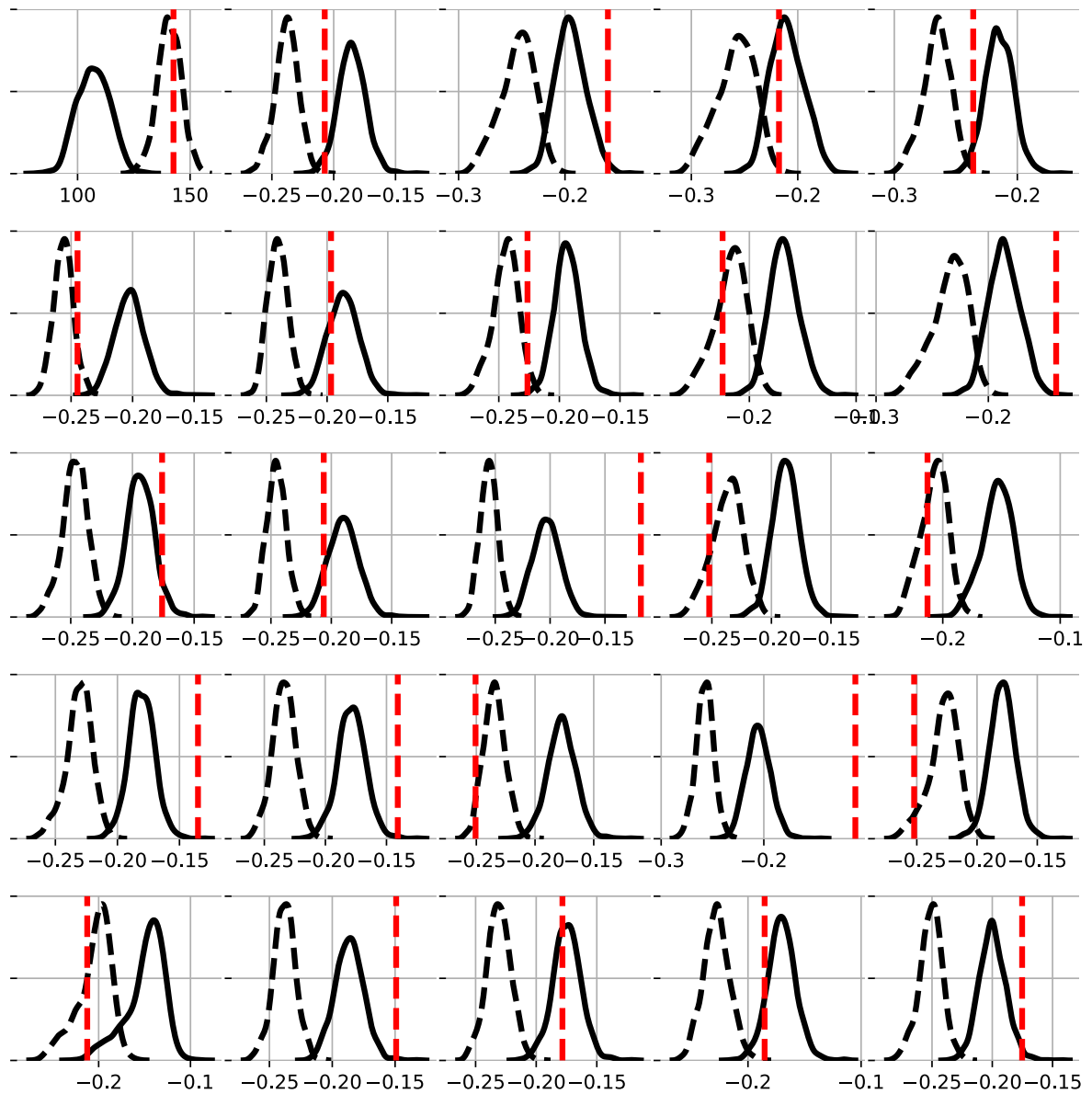
Figure 9.1: Posterior distributions of $\eta(\boldsymbol{\theta})$ for each of the 25 model outputs in two situations: material parameters sampled conditional on volume alone (dashed lines) and material parameters sampled conditional on the measured data. The upper left plot gives results for the LVV and the others show circumferential strains, with each plot containing two different predictive distributions. The solid line is the predictive distribution obtained when inferring the parameters conditional on both volume and circumferential strains while the dashed line is obtained from parameters estimated using only the volume measurement. The vertical dashed lines show the measured value of the output. Notice that the volume is underestimated in order to match the strain measurements.

way: through a departure of the residuals from normality. This means that their influence can be difficult to disentangle. To allow for a proper analysis of discrepancy introduced by the simulator, this chapter will use a series of synthetic data studies to study the effects that different components of the simulator can have on our parameter inference.
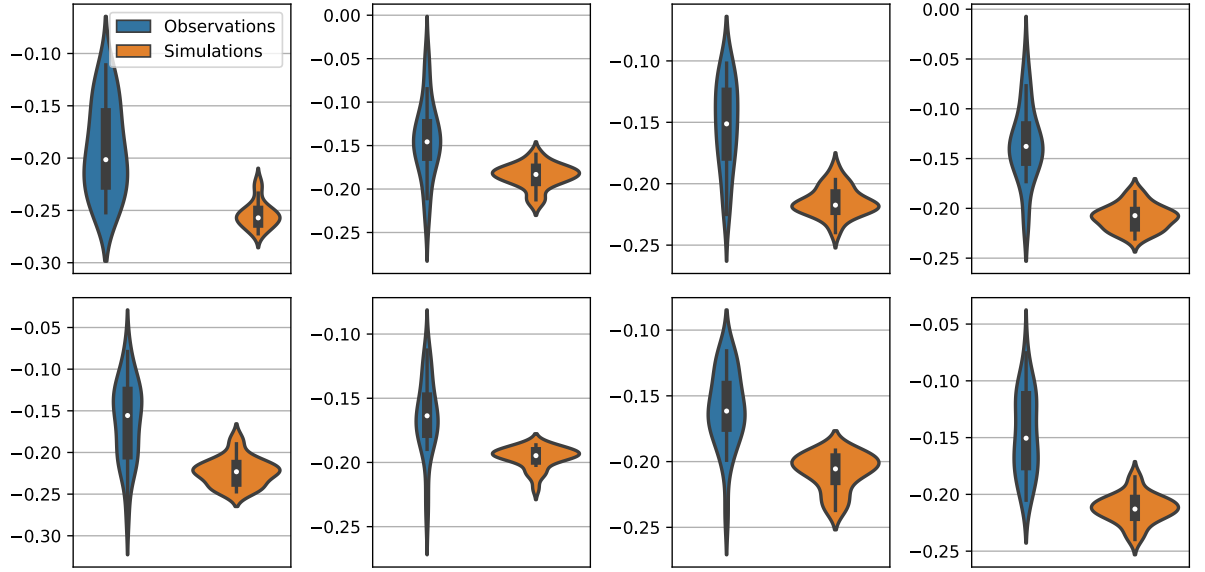
Figure 9.2: Each subplot compares the measured circumferential strain data for a single test case with simulated circumferential strains obtained at parameters that allow us to match the LVV measurement. In all plots the spread of the distributions of measurements is larger than that of the simulations and the median of the simulations is consistenly below the median of the measurements.

9.1.2 The presence of model discrepancy in the LV simulator

For a new subject, with end diastolic pressure (EDP) P (either measured or assumed) and LV geometry $\mathbf{g}$, we may relate a set of measured data $\boldsymbol{y}^*$, obtained from a CMR scan of a subject's LV, with the true underlying strains and standardized LVV, gathered in the term $\boldsymbol{\psi}$:

$$\boldsymbol{y}^* = \boldsymbol{\psi} + \boldsymbol{\xi}, \tag{9.1}$$

When modelling the measured LV data, we approximate the unknown $\boldsymbol{\psi}$ with a simulator, $\zeta(\cdot)$. Similar to the exposition of Goldstein and Rougier (2009), we can consider this simulator as consisting of several parts. The *biomechanical model*, which includes the H-O model and the momentum balance equations, provides the system of equations to be solved during simulation. This is assumed to be fixed, selected by the mathematical modeller during the forward modelling. This model takes as input a set of *constitutive parameters*, $\boldsymbol{\theta}$, and depends on a set of conditions, $\mathbf{u}$, that must be selected in order to evaluate the simulator for a given set of parameters. These *physiological conditions—* which allow us to specify the model for a given subject— include the LV geometry, $\mathbf{g}$, and boundary conditions such as the EDP of the blood within the cavity at end of diastole. Finally, the simulator also requires a *solver*, which consists of the finite element solver and the discretization of the LV geometry into a set of finite element nodes. Like the model,

the solver is fixed by the mathematical modeller and built into the simulator code. The simulator was discussed in greater detail in Section 2.2.

9.1.3 A closer look at the physiological conditions

Accurate model calibration depends on the accuracy with which we can represent the set of physiological conditions for a patient. In principle, these can all be measured in the clinical setting and included in the simulator. In practice, the absence of accurate non-invasive measurement methods, or the vulnerability of existing ones to operator subjectivity, introduces extra uncertainties to the modelling of the LV that decrease the confidence we have in the model and subsequent inferences. These sources of error, which will be explored in simulation studies later in this chapter, will now be discussed in more detail.

Fibre angles

The ventricular myocardium is built up of a stack of fibres, the orientation of which varies as we move from endocardium to epicardium. Advanced imaging techniques, such as diffusion tensor magnetic resonance imaging (DTMRI), have been used to show that the fibres in the myocardium follow a mean fibre direction, subject to some dispersion (random fluctuations around the mean direction) (Guan et al., 2021). This mean orientation can vary from patient to patient and is affected by ventricular remodelling (Palit et al., 2015). In the simulator used for this work, a Laplace Dirichlet rule based approach (Bayer et al., 2012; Wang et al., 2013a) is adopted for the fibre orientation, assuming that the mean fibre angle varies linearly between the endocardium and epicardium. The effect of these bounds was explored in the sensitivity study carried out in Chapter 8, but no consideration was given to how they could impact the inverse problem. Instead, fixed boundaries of $\alpha_{\mathrm{endo}} = -60$ and $\alpha_{\mathrm{epi}} = 90$ were used for the endocardium and epicardium, respectively.

The effect of fibre orientation on the estimation of passive tissue properties has been explored by Palit et al. (2018a), who inferred the passive properties while allowing for different angles in a rule based approach. They found that the fibre angles affected the parameter estimates when optimizing based on an objective function that included the LVV[1]. More recently, Guan et al. (2021) studied the effect of fibre dispersion, which accounts for random perturbation of the fibre orientation around the mean fibre angle. They found that the dispersion especially affects the behaviour in the active contraction part of the cardiac cycle, but still had some effect on the passive filling process. Unfortunately, dispersion is still being explored by the mathematical modellers (Guan et al., 2021) and could not be included in the current simulator code. For this reason, it will not be explored in this work.

---

[1]This objective included LVV and an extra term measuring adherence to the Klotz curve, which was discussed extensively in Chapter 7

Pressure

Easily accounted for in the emulator, as done in Chapter 8, is the EDP of the blood within the LV cavity. Problematic, however, is the invasiveness of measurement methods, which means that we can only rely on measured values when dealing with diseased patients. For healthy volunteers, a value is usually assumed, which in most of the work in this thesis has meant setting the EDP to a value of 8 mmHg. This value, motivated by Bouchard et al. (1971), inevitably introduces some bias to the parameter estimation. However, the work of Palit et al. (2018a) showed that the parameter estimates were fairly robust when varying pressure around values that we expect in healthy volunteers (between 8 and 12 mmHg).

LV geometry

Calibrating the LV simulator to a particular patient depends, importantly, on the LV geometry. Unlike pressure and fibre angles, non-invasive methods currently exist for measuring the LV geometry in vivo, using the procedures outlined in Section 5.2.1. In previous chapters, a PCA representation of the LV geometry was adopted, while acknowledging in Chapter 6 that the failure of this method to take the biomechanical model into account makes it suboptimal for LV geometry representation in the simulator. This idea of the LV geometry representation being suboptimal was built on the premise that the LV geometry exracted from a CMR scan can be treated as a ground truth. However, in reality this makes little sense because the final shape extracted from the CMR scan will be subject to both inter and intra observer (operator) errors. As a result, the effect that these different operators have on the biomechanical model is of interest for two reasons. Firstly, the variation introduced to the inferred cardio-mechanics as a result of operators is of interest. Secondly, the utility of a PCA basis for representing the LV geometry in the simulator depends critically on its comparison with the performance of different operators: if LV modelling errors incurred using a PCA representation of the LV geometry are lower than those incurred with a new operator then PCA is a sufficient representation of the LV geometry. At the time of this work the effect of operators on the mechanical modelling of the LV had not been explored, but recently this was carried out by Hasaballa et al. (2021) who found that the operator could have a substantial effect on the results of the parameter estimation. However, they did not benchmark this operator effect against the performance of PCA. In this chapter, it will be shown that errors introduced to parameter estimation as a result of PCA are in line with the error introduced by a new operator.

9.1.4   Effect of erroneous physiological conditions and further modelling errors

It is worth considering the consequences of erroneous physiological conditions on the noise model from (9.1). Assuming that the simulator has the capacity to perfectly replicate the underlying behaviour of the real world, we can take the ground truth constitutive parameters, $\tilde{\boldsymbol{\theta}}$, and physiological conditions, $\tilde{\mathbf{u}}$, to be those that provide the perfect fit to the true underlying process. The simulator can be incorporated into (9.1) by replacing $\boldsymbol{\psi}$ with ζ evaluated at this parameter combination:

$$\boldsymbol{y}^* = \zeta(\tilde{\boldsymbol{\theta}}, \tilde{\mathbf{u}}) + \boldsymbol{\xi}, \tag{9.2}$$

Note that the LV geometry $\mathbf{g}$ has been absorbed by $\tilde{\mathbf{u}}$ because the LV geometry is also considered as a physiological condition. The noise model is the same as before due to the perfect correspondence between simulator and reality. In any use of the simulator model, we rely on measured or assumed values for $\mathbf{u}$, which we may represent as $\hat{\mathbf{u}}$. When these are measured, such as the case of the LV geometry, they will be subject to measurement error (or possibly approximation error in the case of the statistical emulator with PCA) and when they are assumed known and fixed to population based values, they will be subject to some bias. In the presence of these suboptimal modelling conditions, the best fitting constitutive parameters will no longer be the same as the "ground truth" values. At the new best fitting parameters $\hat{\boldsymbol{\theta}}$, we could argue that it no longer makes sense to think of the noise model the same as before, instead considering the following model:

$$\boldsymbol{y}^* = \zeta(\hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}) + \boldsymbol{\xi}^* \tag{9.3}$$

where $\boldsymbol{\xi}^*$ will often not have the same variance and may no longer satisfy the independence assumption. For instance, in the presence of incorrect fibre angles we might expect a relationship between the error in two neighbouring regions of the model. Critically, if we were to assume the same noise model as was considered in (9.2), then a bias would be introduced to the estimated parameters.

It turns out that the inconsistencies between the model and reality are more fundamental than erroneous physiological conditons. Built into all parts of the simulator are several simplifying assumptions and chief among these is the model itself. The H-O model is not a fundamental law of nature to which the myocardium abides. Instead, it is a mathematical description of the myocardium, constructed by modellers and validated against experimental data to prove its ability to match the behaviour of the myocardium tissue at different ranges of stretch (Holzapfel and Ogden, 2009). As such, we can be confident that, although accurate, this description of the tissue will not perfectly replicate the en-

ergy build-up in the myocardium. A further inconsistency is in our use of this law in the simulator, where we make the assumption that the myocardium tissue can be described by homogenous material properties. Even for healthy tissues we would expect this assumption is violated, but any introduction of heterogeneity to the model is highly complex and methods do not currently exist to allow for it in the simulator. This reality of heterogeneity does introduce the question of how we can justify an interpretation of homogenous tissue properties, especially if we do not account for any resulting mismatch in the simulator. A final limiting assumption is that the model considers the LV in isolation. In reality, the LV motion at the septum (see Figure 2.3) will be limited by the presence of the right ventricle (RV). Using a bi-ventricular model would better resemble the boundary conditions on the wall of the LV, without which we will incur a discrepancy between the reality and the simulator, as studied by Palit et al. (2015).

9.2 Methodology

In Figures 9.1 and 9.2, we saw evidence of systematic mismatch between the simulator and reality. The possible sources of this discrepancy, specifically related to the mathematical model, were discussed in Section 9.1.2. This section will present an approach that allow us to acknowledge this discrepancy in our noise model, largely based on the methods for parameter calibration in the presence of model discrepancy that were discussed in Section 3.2.7.

9.2.1 Acknowledging model discrepancy in Bayesian model calibration

Consider the situation where we wish to infer the material parameters from measured data $\boldsymbol{y}^*$, with an LV geometry vector $\mathbf{g}$ that has a 5 PC projection $\mathbf{h}$. In order to acknowledge model discrepancy, a likelihood of the following form will be considered:

$$p(\boldsymbol{y}^*|\boldsymbol{\theta}) = l\mathrm{V}(\boldsymbol{\theta}) - \frac{k}{2}\log(\sigma_\mathrm{S}^2) - \frac{1}{2\sigma_\mathrm{S}^2}\sum_{i=1}^{k}(y_\mathrm{i}^* - (\eta_i(\boldsymbol{\theta}, \mathbf{h}) + b_i))^2, \tag{9.4}$$

where $\eta_i(\boldsymbol{\theta}, \mathbf{h})$ is the emulator prediction for output i at parameters $\boldsymbol{\theta}$ and 5 PC representation $\mathbf{h}$. The first term, which is the log-likelihood term associated with the volume measurement, has the following form:

$$l\mathrm{V}(\boldsymbol{\theta}) = -\frac{1}{2}\log(\sigma_V^2) - \frac{1}{2\sigma_V^2}(y_0^* - \eta_0(\boldsymbol{\theta}, \mathbf{h}))^2, \tag{9.5}$$

and σ_{S}^2 is the variance associated with the strain measurements. The volume variance σ_V^2, which is non-identifiable given only one volume measurement, will be fixed during inference to a value that is consistent with an operator error study (see Table E.1) while the strain variance will be inferred along with the constitutive parameters. Within the likelihood in (9.4), an extra term b_i has been included. This extra parameter, which must be inferred along with the constitutive parameters, adds an offset to the prediction for strain i, making the strain more consistent with the volume. To summarize: the model in (9.4) assumes that there is no systematic mismatch between the predicted and measured LVV, which is necessary because of the absence of prior knowledge about the model discrepancy. Moreover, in the presence of only one volume measurement we cannot expect to learn a model discrepancy term on this quantity.

9.2.2   GP model for circumferential strain discrepancy

We have 24 circumferential strains, each of which will be subject to a separate bias correction. Modelling the corrections independently would lead to overfitting and a model that is non-identifiable. To overcome this, we must model some dependence between the correction variables, encouraging smoothness that will help to limit the flexibility of the function. Following the work of Brynjarsdóttir and O'Hagan (2014), this can be done by introducing a GP prior on the discrepancy function, allowing for the correction to be jointly sampled with the constitutive parameters according to the following likelihood function:

$$p(\boldsymbol{y}^*|\boldsymbol{\theta}) = l\mathrm{V}(\boldsymbol{\theta}) - \frac{k}{2}\log(\sigma_{\mathrm{S}}^2) - \frac{1}{2\sigma_{\mathrm{S}}^2}\sum_{i=1}^{k}(y_{\mathrm{i}}^* - (\eta_i(\boldsymbol{\theta},\mathbf{h}) + b_i(\mathbf{x}) + v))^2, \qquad (9.6)$$

where the parameter v is needed due to a zero-mean assumption on the GP. Assigning a GP prior and treating $\mathbf{b}$ as a function, we require a domain $\mathcal{X}$ that defines the space over which the discrepancy varies. On this domain we can then specify a kernel function that provides the covariance between realizations of $\mathbf{b}$ at two separate points, as required for the GP prior. For identifiability purposes, we require that $\mathcal{X}$ is independent of $\boldsymbol{\theta}$.

To account for model discrepancy in the LV problem, the domain of $\mathbf{b}$ is assumed to be the strain location. The location of the strains on the LV wall was visualized in Figure 2.3, where the coloring indicated a different circumferential positioning. One possible choice for $\mathcal{X}$ is to consider the strain values defined over 1D space, with an ordering that accounts for their location on the LV surface. However, any ordering we consider here will be suboptimal because we cannot fully define the positioning of 3D strain locations in only 1 dimension. Instead, we could use 3D coordinates in Euclidean space to define the positioning of the strain locations and use a standard GP kernel, like

the squared exponential, to define the correlations between locations. The issue with such an approach is that the distance measure can jump across the centre of the cavity and misrepresent the distance between two strain locations. Ideally, we want to account for the circular shape of each short axis slice, for which we require a GP kernel defined on polar coordinates.

Let us assume the strain measurements lie on a grid in two dimensional cylindrical coordinates with a circumferential coordinate from $\left\{0, \frac{\pi}{3}, \frac{2\pi}{3}, \pi, \frac{4\pi}{3}, \frac{5\pi}{3}\right\}$ and z coordinate from $\{0, 1, 2, 3\}$. Thinking first about how to define the correlation in the circumferential direction, we can consider the geodesic distance between two regions, which is given by $d(\gamma, \gamma') = \mathrm{acos}(\cos(\gamma - \gamma'))$ where γ and γ' are the circumferential positioning of the two regions of interest. Naturally, we can apply the polar GP kernel that was discussed in Section 3.4.2 to specify correlations in this domain. In the z-direction, the strain locations are defined in Euclidean space and we can use standard kernel functions like squared exponential or Matérn class kernels. Combining this kernel in the z-direction and the polar kernel in the circumferential direction provides us with a kernel function defined on the surface of a cylinder. These can be combined additively or multiplicatively if we wish to account for interactions between the two dimensions. Multiplying two kernels can be thought of as an AND operation. That is, if you multiply together two kernels, then the resulting kernel will have high value only if both of the two base kernels have a high value. Adding two kernels can be thought of as an OR operation. That is, if you add together two kernels, then the resulting kernel will have high value if either of the two base kernels have a high value. In this work, an additive structure will be assumed as it was found to lead to smoother corrective functions in preliminary runs. The final kernel has the following form:

$$k(\{\gamma, z\}, \{\gamma', z'\}) = \sigma_{f_1}^2 W_c(d(\gamma, \gamma'); c, \tau) + \sigma_{f_2}^2 k_z(z, z'; \lambda) \tag{9.7}$$

with the hyperparameters of the model summarized in Table 9.1.

9.2.3   Inference for the model with discrepancy correction

Inference will be carried out using MCMC. Specifically, a HMC within Gibbs sampler will be adopted, where sampling of the hyperparameters is performed adjacent to sampling of the main model parameters. Table 9.1 provides a summary of the parameters present in the model along with the corresponding prior that is used during inference. The bounds for the uniform priors on $\boldsymbol{\theta}$ were as follows: $a, a_\mathrm{f} \in [0.05, 10]$ and $b, b_\mathrm{f} \in [0.1, 30]$.

The posterior of the discrepancy function $\mathbf{b}$ is highly correlated due to the multivariate normal prior that results from the assumption of a GP. To improve efficiency of sampling,

Table 9.1: Description of all parameters from the model in (9.6) along with priors used in Bayesian inference. IG represents an inverse Gamma distribution.

Parameter	Description	Prior ($\pi(\cdot)$)	Rationale
σ_S^2	Noise variance of strains	IG(0.001,0.001)	Conjugate and vague
$\boldsymbol{\theta}$	Material parameters $(a, b, a_{\mathrm{f}}, b_{\mathrm{f}})$	Uniform	Physiological bounds
v	Global bias of strains	Normal(0.05)	Knowledge of small global discrepancy
λ	Lengthscale of z-component of kernel function	Gamma(2,2)	Encourage smoother functions
$\sigma_{f_1}^2$	Signal variance of polar kernel	IG(0.001,0.001)	Vague
$\sigma_{f_2}^2$	Signal variance of z-direction kernel	IG(0.001,0.001)	Vague
c	Compact support parameter of polar kernel	$u(0, \pi)$	Physical bound
τ	Smoothness parameter of polar kernel	$u(4, 10)$	Encourage smoother functions

this is sampled in the whitened space:

$$\tilde{\mathbf{b}} = \mathbf{L}^{-1}\mathbf{b} \tag{9.8}$$

where $\mathbf{b} \sim \mathrm{GP}(\mathbf{0}, \mathbf{K})$ and $\mathbf{K} = \mathbf{L}\mathbf{L}^\top$ is the Cholesky decomposition. This method was used by Kuss and Rasmussen (2005) to improve sampling efficiency, and depends on the selection of the hyperparameters to determine the matrix $\mathbf{L}$. Due to the Gibbs step, we only compute the Cholesky decomposition once per MCMC sample because the GP hyperparameters have been sampled before the model parameters. The procedure is outlined in Algorithm 4.

Algorithm 4 HMC within Gibbs for the polar GP model.

1: Sample from $p(\sigma_S^2|\boldsymbol{\theta}, v, \lambda, \sigma_f^2, c, \tau, \mathbf{b}, \boldsymbol{y}^*)$, which is possible in closed form given conjugate inverse Gamma prior.
2: Sample from $p(\lambda, c, \tau, \sigma_{f_1}^2, \sigma_{f_2}^2|\boldsymbol{y}^*, \boldsymbol{\theta}, \sigma^2, \mathbf{b}, v)$ using HMC.
3: Calculate updated covariance matrix $\mathbf{K}$ and corresponding cholesky decomposition: $\mathbf{K} = \mathbf{L}\mathbf{L}^\top$.
4: Sample from $p(\boldsymbol{\theta}, \mathbf{b}, v|\boldsymbol{y}^*, \sigma_S^2, \lambda, c, \tau, \sigma_{f_1}^2, \sigma_{f_2}^2)$ using HMC, where $\mathbf{b}$ is sampled in whitened space.

The posterior distributions used in steps 2 and 3 of Algorithm 4 have the following

form:

$$p(\lambda, c, \tau, \sigma_f^2 | \boldsymbol{y}^*, \boldsymbol{\theta}. \sigma_S^2, \mathbf{b}, v) \propto \mathcal{MVN}(\mathbf{b}|0, \mathbf{K}(\lambda, c, \tau, \sigma_{f_1}^2, \sigma_{f_2}^2)) \pi(\lambda) \pi(c) \pi(\tau) \pi(\sigma_f^2)$$

$$p(\boldsymbol{\theta}, \mathbf{b}, v | \boldsymbol{y}^*, \sigma_S^2, \lambda, c, \tau, \sigma_{f_1}^2, \sigma_{f_2}^2) \propto p(\boldsymbol{y}^* | \boldsymbol{\theta}, \mathbf{b}, v) \mathcal{MVN}(\mathbf{b}|0, \mathbf{K}(\lambda, c, \tau, \sigma_{f_1}^2, \sigma_{f_2}^2)) \pi(v)$$

where the various prior distributions, represented by $\pi(\cdot)$, are found in Table 9.1 and $\mathbf{K}(\lambda, c, \tau, \sigma_{f_1}^2, \sigma_{f_2}^2)$ is the GP covariance matrix.

9.2.4   The emulator model

Based on the results from Chapter 8, a four dimensional parameterization of the H-O model is adopted where a, b, a_f, b_f are considered as random variables while a_s, b_s, a_{fs}, b_{fs} are fixed to population based values from the literature (Gao et al., 2017). To allow for variations in the LV geometry, a 5PC representation of the LV geometry is included in the input space of the emulator, resulting in a 9 dimensional regression problem. The same simulator outputs as previous chapters were used: 24 circumferential strains and the LVV. Essentially, this emulator is the same as the one from Chapter 6 but a different parameterization of the H-O model is adopted.

In Chapter 6, a Sobol sequence was used to generate the training data for the emulator, despite the evidence from Chapter 4 that the function is nonstationary. In Chapter 8, it was shown that a warping of input space provided by a log transformation performed similarly to more flexible neural network transformations of the input space, motivating the use of a Sobol sequence on log-uniform space. This method will be used to generate the training data for the current emulator. In total, 30,000 simulations were obtained for training the emulator model and, as in Chapter 6, a neural network was used. To decide on a final network, different combinations of 50,100,120 and 200 nodes were used in two hidden layers, each containing tanh activation functions. The optimal architecture was selected based on the MSE on a validation set of 3000 simulations. Training was performed using the Adam optimizer with a decaying learning rate starting at 0.01. As with other neural network models in this thesis, the validation set also allowed for experimentation with different batch sizes. Based on previous results, the weights were initialized using samples from a $\mathcal{N}(0, 0.1)$ distribution.

9.3   Implementation

Section 9.1.2 discussed the idea of discrepancy as a result of physiological conditions. These included the LV geometry, the fibre orientation and the EDP. In order to explore the effects of these on our model calibration, we can generate synthetic data by running the simulator at a series of different physiological conditions. Several simulation studies

to determine the effect of physiological conditions on the model calibration will now be outlined.

9.3.1 Joint optimization of fibre angles and constitutive parameters

The data in Figure 9.1 are subject to a substantial degreee of noise. Despite this, we could see that the model and reality are in disagreement: in order to match the volume, most of the circumferential strains must be overestimated. As we learnt in Section 9.1.2, there are several sources of error in the model related to assumptions being made in the physiological conditions and the typical approach taken in this thesis was to fix the physiological conditions to sensible values based on knowledge from the literature. In light of this, it is worth questioning whether we can get closer to the measured data by estimating the physiological conditions from the data. Not all conditions can be considered here: due to its high-dimensional representation the LV geometry is difficult to tune to the data, and jointly inferring EDP and material parameters is an ill-posed problem. On the other hand, the fibre angles can be represented in a two dimensional space and, until now, their effect on the parameter estimation has not been thoroughly explored. In this study, the angles will be jointly optimized with the material parameters using Bayesian optimization (BO) (Shahriari et al., 2016), which has been used to optimize the material parameters by Romaszko et al. (2021). In doing so, we treat the angles as tuning parameters, and their value is only a secondary concern. Primarily, we are hypothesizing that the fixing of the fibre angles represents a primary source of discrepancy between the model and reality and wish to see whether treating these parameters as tunable can help us to better match the LVV and circumferential strains. This will be investigated for each of the 8 measured test sets.

The general idea in BO is to replace the iterative optimization of an objective function $f(\boldsymbol{\theta})$, with iterative optimization of an acquisition function $A(\boldsymbol{\theta})$. Starting with a set of objective function evaluations $\{\boldsymbol{\theta}_i, f(\boldsymbol{\theta}_i)\}_{i=1}^{n}$, a GP (or similar regression model) is fitted to approximate f. Using information from the fitted GP, an acquisition function is defined that will be optimized to select a new query point at which to evaluate the objective function. This acquisition function typically has a multi-modal landscape, but is very computationally cheap to evaluate and optimize. Once the query point is added to the training set $\{\boldsymbol{\theta}_i, f(\boldsymbol{\theta}_i)\}_{i=1}^{n+1}$, the GP is updated. Updating the GP that is fitted to the objective function evaluations also updates the acquisition function, which can then be re-optimized to find the next query point. Importantly, the form of $A(\boldsymbol{\theta})$ is such that there is a trade-off between exploring unseen locations in input space, and exploiting areas where we think the objective function is low (assuming minimization). The fact that the acquisition function encourages exploration as well as exploitation means that the crudeness of the

objective function approximation, especially at early stages, is not detrimental. As long as the posterior distribution of the function correctly represents our state of belief, we will explore areas where the uncertainty is high and exploit areas where the posterior mean is low. In particular, the algorithm will initially explore the input space, before exploiting as the fitted GP becomes more confident. Further details can be found in the literature (Shahriari et al., 2016).

To use BO in this study, an initial set of 300 points (from a LHD) were run in parallel before initializing the BO algorithm. The benefit of such a large number of initial runs is that we can easily exploit parallelization at this stage of the algorithm. This LHD gives a matrix of design points $\mathbf{X}$ that is $300 \times p$, where p is 4 for the fixed physiological condition runs (only material parameter space) and 6 for the optimized angles runs (material parameters plus two angles). To get the initial design for BO $\tilde{\mathbf{X}}$, the first 4 dimensions of $\mathbf{X}$ are transformed as follows:

$$\tilde{\mathbf{x}}_i = \exp(\mathbf{x}_i)$$

and rescaled to $[0.1, 10]$, which was the physiological range of the parameters in this study. As we see in Figure 9.3, this newly transformed design (labelled as the exp design) is a compromise between the highly skewed log-uniform design and the uniform design.
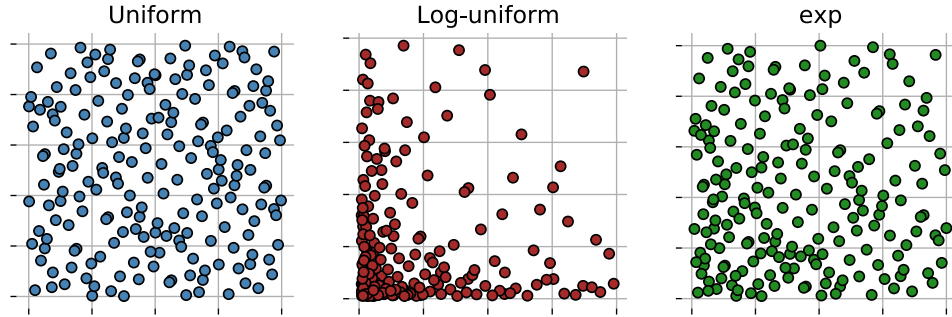

Figure 9.3: Comparing the different designs. The log-uniform design is better for building an emulator because most points are in the highly varying portion of the function, but many of these parameter values will not be physiologically realistic. The exp design used in this chapter is a compromise between the standard Sobol sequence and the Sobol sequence in log-Uniform space.

From the initial points, a set of simulations $\mathbf{Y}$ was obtained, and the LVV was standardized using the early diastole volume. The mean and standard deviation of these simulations (one for each of the 25 outputs) were used to standardize the initial data and the test data. The residual sum of squares was used as the objective function, but preliminary runs suggested that with volume and strains equally weighted there was not enough incentive to match the volume. Therefore, the contribution of the LVV to the RSS was

upweighted by a factor of 2, ensuring that the optimizer tries to match the LVV as well as the circumferential strains. The BO was run for 300 iterations after the initial set of data were generated, with the lower confidence bound acquisition function (Shahriari et al., 2016) used to propose the next query point in the BO algorithm. In the LCB acquisition function, the default weighting between mean and standard deviation of the GP predictive distribution in the MATLAB BO toolbox was used:

$$A(\boldsymbol{\theta}) = \mu(\boldsymbol{\theta}) - 2\sigma(\boldsymbol{\theta}) \tag{9.9}$$

This acquisition function promotes exploration, which makes sense when we are in such a high dimensional space.

9.3.2   Accounting for fibre angle discrepancy

To explore the effect of fibre angle approximation, the emulator from Section 9.2.4 will be used to infer the material parameters. Recall that the emulator was trained on simulations obtained from different combinations of constitutive parameters and 5PC representations of the LV geometry, with the fibre angles held fixed at $\alpha_{\text{endo}} = -60$ and $\alpha_{\text{epi}} = 60$. These are standard values of the fibre angles, used in the work in the earlier chapters of this thesis and found in the literature (Noè et al., 2019). To introduce discrepancy between the model and the test data as a result of fibre angle discrepancy, test data were generated with fibre angles and constitutive parameters set to combinations of the following values:

$$\begin{aligned} (\alpha_{\text{endo}}, \alpha_{\text{endo}}) &= \{(-60, 90), (-30, 60), (0, 90)\} \\ (a, b, a_{\text{f}}, b_{\text{f}}) &= \{(0.1, 5, 1, 6), (0.5, 8, 1.5, 6), (1, 1, 1, 1), (1, 5, 1, 5)\} \end{aligned} \tag{9.10}$$

These material parameter vectors were selected to give physlogically realistic diastolic behaviour. Each combination of material parameters and fibre angles was used, giving 12 synthetic test sets in total. Gaussian noise with standard deviation equal to 5 ml was added to the volume (motivated by the study in Table E.1) and iid Gaussian noise with a standard deviation of 0.03 was added to the circumferential strains (as discussed in Chapter 8, this is consistent with results from the literature). In order to remove the effect of LV geometry error on the fibre angle exploration, the simulated data were generated from a synthetic LV geometry generated in the 5 PC space.

Of interest is the effect that erroneous fibre angles can have on the estimation of the material parameters. In addition to this, we can investigate the use of methods from Section 9.2.2 to correct for the discrepancy introduced by the erroneous fibre angles. Given the nature of the discrepancy model, we might expect that inference is driven by the LVV component of the model. This concern will be addressed by comparing with inference results based on LVV alone. For each model, inference will be performed using MCMC.

For the volume only model and the model without discrepancy correction, this is a simple application of NUTS in line with the procedures used throughout this thesis. In the case of the discrepancy correction model, the procedure outlined in Section 9.2.3 will be used. Convergence is assessed using the PSRF, with 1000 samples obtained after convergence.

9.3.3 Effect of LV geometry

In order to infer the material parameters from measured data, we require the LV geometry. During parameter inference, we fix the LV geometry as an input to the simulator and optimize or sample the constitutive parameters. In the real data case, the measured data are obtained from CMR scans with the LVV obtained by reconstruction of the LV geometry at end of diastole (this follows the same procedure as segmentation of the early diastole LV geometry, which was outlined in Section 5.2.1) and the circumferential strains obtained using an independent measurement procedure called deformation tracking (Mangion et al., 2016a). Ideally, the effect of operator error would be assessed using data measured from CMR scans. For multiple different subjects, we could get multiple operators to segment the LV geometry at early and end of diastole and then measure the circumferential strains from the CMR scans. Unfortunately, the inconsistencies between model and data that we witnessed in Section 9.1.1 make this a difficult analysis to perform. Moreover, the difficulty in measuring circumferential strains means that the errors in their measurement would overshadow the errors in the LV geometry representation, which is of principal interest here. As a result, a different study had to be constructed.

Of interest is the effect that different operators (different people segmenting the LV geometry) have on the biomechanical model. In total, the LVs of 6 different patients[2] were segmented by three different operators who will be referred to as operators 1,2 and 3. The operators had different levels of experience in segmentation of the LV, but all were trained for the task. To investigate the effect of the operator on parameter estimation, the procedure visualized in Figure 9.4 was used.

Each operator will be treated as a *main operator*, with the other two considered as *test operators*. The main operator's LV geometry G, which has vector representation $\mathbf{g}$, will be used to generate a set of synthetic data by evaluating the simulator using $\mathbf{g}$ and material parameters $\tilde{\boldsymbol{\theta}} = (a, b, a_\mathrm{f}, b_\mathrm{f}) = (0.60, 1.30, 1.00, 3.60)$[3]. This is repeated for each of the 6 different patients, creating 6 "measured" datasets for each main operator. To infer the parameters from these datasets, several test geometries $\hat{G}$, which have vector representation $\hat{\mathbf{g}}$, will be used. These are the LV segmentations of the test operators, as well as the 5 and 10 PC reconstructions of the main operator's segmented LV geometry. If

---

[2]These were obtained from a study of COVID patients.

[3]The material parameters were selected because the simulated LVV corresponds to an EF roughly equal to 60%, which is a normal value (Hudson and Pettit, 2020).

the error introduced to the parameter estimates by PCA is lower than the error introduced by a new operator, then we have evidence that the inclusion of any extra LV geometry information will not improve the parameter inference accuracy, given the current methods for measuring the LV geometry.
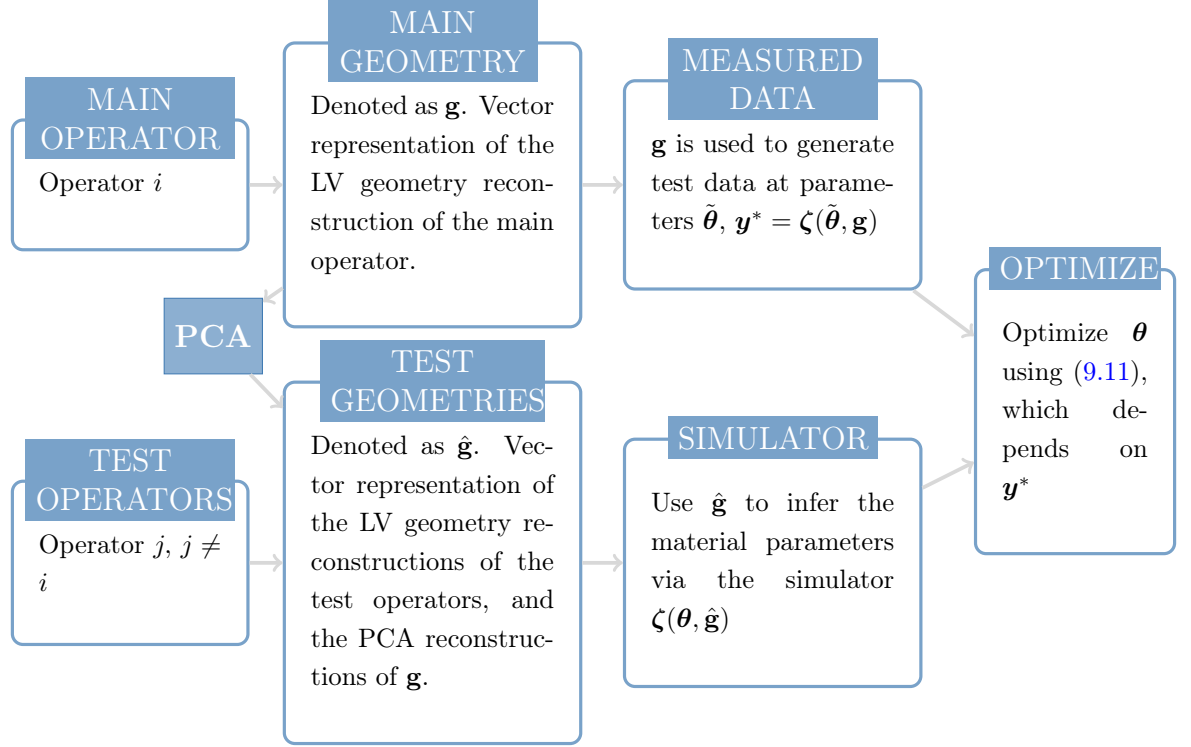

Figure 9.4: Diagram of the optimization procedure for the LV geometry study. In total there are 4 test LV geometries, coming from the two test operators and the 5 and 10 PC reconstruction of the main geometry. Therefore, the optimization is repeated 4 times.

For the following outline of the optimization procedure, let $\boldsymbol{y}^* = \zeta(\tilde{\boldsymbol{\theta}}, \mathbf{g})$ be the measured data. The first output, y_0^*, is the LVV and the remaining 24, $\{y_i^*\}_{i=1}^{24}$, are the measured circumferential strains. To optimize the material parameters, the following objective function is used:

$$L(\boldsymbol{\theta}) = \left(\frac{V_{\text{early}} - y_0^*}{y_0^*} - \frac{\hat{V}_{\text{early}} - \zeta_0(\boldsymbol{\theta}, \hat{\mathbf{g}})}{\zeta_0(\boldsymbol{\theta}, \hat{\mathbf{g}})} \right)^2 + \sum_{i=1}^{24} (y_i^* - \zeta_i(\boldsymbol{\theta}, \hat{\mathbf{g}}))^2, \qquad (9.11)$$

where V_{early} is the volume of G at early diastole and $\hat{V}_{\text{early}}$ is the volume of $\hat{G}$ at early diastole (this is in contrast to y_0^*, which is the LVV). The volume component of this loss function is unconventional and requires some discussion. The motivation behind this term is to negate the effect that the initial volume of a reconstruction has on the parameter optimization, instead focusing on the effect of the shape of the LV geometry. This does not remove any contribution of LVV from the loss function, instead it means that the contribution from LVV is not dominated by the difference in early diastole volume. The

optimization was performed in MATLAB using *fmincon* with the sequential quadratic programming method, which was also used by Gao et al. (2015). Each optimization is initialized from the ground truth material parameter value, so the optimizer measures the perturbation introduced as a result of the LV geometry approximation.

9.4 Results

9.4.1 An initial look at the effect of fibre angles

In Figure 9.5, simulations obtained using the same LV geometry and material parameters but with a range of different fibre angles are plotted. The line plots show the trend in the circumferential strains, which can vary quite substantially as the angle changes. The corresponding angles are provided in the legend along with the LVV (in the legend it is labelled as EDV). The effect of the angle change is similar no matter the LV geometry, which is that some circumferential strains are affected quite substantially, especially the anterior lateral and inferior lateral regions. An important observation to take from this plot is as follows: by varying the angle from $\alpha_{\mathrm{endo}} = -60, \alpha_{\mathrm{epi}} = 90$ to $\alpha_{\mathrm{endo}} = -30, \alpha_{\mathrm{epi}} = 60$, we get only a 2 ml difference in LVV but a more substantial increase in the circumferential strains, particularly in the anterior lateral, inferior lateral and inferior region. In other words, it appears that the effect of fibre angle on volume and circumferential strain is not always in the same direction. This is precisely the type of effect that could be missing when we calibrate the simulator to new data, as shown in Figure 9.1.
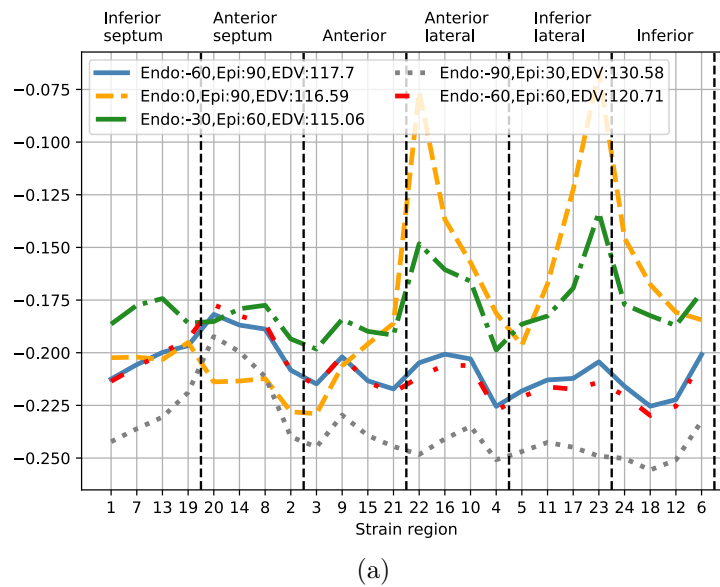

(a)

Figure 9.5: Effect of fibre angle on the simulated circumferential strains and LVV.

9.4.2 Bayesian optimization of angles and material parameters

The second exploration of the fibre angles considers the task of learning the material parameters and fibre angles together from data measured from CMR scans. Using BO, the four constitutive parameters a, b, a_f, b_f were optimized, with final results provided in Table 9.2. As well as providing the optimal configuration of the constitutive parameters, the table provides the error in the LVV at the optimum as well as the error in the circumferential strains, assessed as the mean absolute error in the 24 circumferential strains. We see a common inability to fit both the circumferential strains and the LVV. Both the angles in the table are constant because these were held fixed during optimization. Table 9.3 gives the results of the 6 dimensional BO problem, including the four constitutive parameters and the two fibre angles from the RBM. Comparing the volume errors, we see that by optimizing the fibre angles along with the constitutive parameters we can better match the volume measurements while still reaching the same (or lower) errors in the circumferential strains. Typically, this requires that the two angles are quite small compared with the values used to obtain the results in Table 9.2 and, as discussed above, these were treated as tuning parameters with no attention paid to their physiological validity.

Table 9.2: Results of Bayesian optimization of material parameters, with fibre angles fixed at $\alpha_{\text{endo}} = -60$ and $\alpha_{\text{epi}} = 90$.

							Fixed angles Vweight=2	
ID	a	b	a_f	b_f	α_{endo}	α_{epi}	Absolute LVV error	Median absolute strain error
HV1	0.30	0.54	0.60	9.74	-60	90	20.58	5.05×10^{-2}
HV2	0.12	9.91	8.32	0.92	-60	90	12.48	4.0×10^{-2}
HV3	0.10	5.88	3.17	9.94	-60	90	23.8	3.67×10^{-2}
HV4	0.18	9.80	2.68	10.0	-60	90	35.50	4.18×10^{-2}
HV5	0.10	6.88	1.98	9.96	-60	90	17.2	4.71×10^{-2}
HV6	2.35	0.27	3.83	1.59	-60	90	16.0	2.62×10^{-2}
HV7	0.14	9.90	1.29	9.8	-60	90	15.0	3.08×10^{-2}
HV8	0.12	3.18	3.76	9.98	-60	90	29.8	4.24×10^{-2}

9.4.3 An attempt to account for fibre angle discrepancy

Based on the results of the BO study, it appears that fixing of fibre angles introduces a source of model discrepancy. To explore how this discrepancy can be acknowledged during parameter inference, the model from Section 9.2.3 will now be considered, inferring the parameters from the test data that were generated according to Section 9.3.2. Comparisons in parameter space are confounded by the strong correlations between the parameters of the H-O model, so another method for comparison is required. Recall from (9.10) that the test data were generated using each pair of material parameter-fibre angle pairs from

Table 9.3: Results of Bayesian optimization of fibre angles and material parameters.

							Optimized angles Vweight=2	
ID	a	b	a_{f}	b_{f}	α_{endo}	α_{epi}	Absolute LVV error	Median absolute strain error
HV1	0.36	1.27	3.00	0.58	-11.2	24.8	1.23	4.48×10^{-2}
HV2	0.12	9.95	5.84	1.01	-33.7	89.7	3.13	3.93×10^{-2}
HV3	0.18	3.52	5.01	9.91	-17.7	5.5	1.11	3.79×10^{-2}
HV4	0.12	5.55	4.63	6.26	-36.2	25.9	16.78	3.82×10^{-2}
HV5	0.16	5.79	3.37	0.96	-1.9	36.2	8.7	4.13×10^{-2}
HV6	1.43	3.67	2.19	0.54	-33.0	52.7	8.94	2.64×10^{-2}
HV7	0.15	4.31	3.03	5.19	-7.3	5.8	3.12	2.86×10^{-2}
HV8	1.69	0.35	3.65	2.90	-2.73	19.3	12.5	3.76×10^{-2}

sets of 4 material parameters configurations and 3 fibre angle configurations. Letting $\tilde{\boldsymbol{\alpha}}$ be a particular fibre angle configuration from this set of 3 and $\tilde{\boldsymbol{\theta}}$ be a material parameter vector from the set of 4 different parameter configurations, then the test data are:

$$\boldsymbol{y}^* = \boldsymbol{\zeta}(\tilde{\boldsymbol{\theta}}, \tilde{\boldsymbol{\alpha}}, \mathbf{g}) + \boldsymbol{\xi} = \mathbf{z} + \boldsymbol{\xi} \tag{9.12}$$

where $\mathbf{g}$ is the representation of an LV geometry generated from a point, $\mathbf{h}$, in 5 PC space. The variable $\boldsymbol{\xi}$ corresponds to the Gaussian noise, with standard deviation of 5 ml for the volume and 0.03 for circumferential strains. If we were to plug $\tilde{\boldsymbol{\theta}}$ and $\mathbf{h}$ into the emulator η, we would recover:

$$\tilde{\boldsymbol{y}} = \boldsymbol{\eta}(\tilde{\boldsymbol{\theta}}, \mathbf{h}) \neq \mathbf{z} \tag{9.13}$$

where the lack of equality stems from the fact that the emulator was trained on simulations that assume a fixed fibre angle configuration of $\alpha_{\mathrm{endo}} = -60°$, $\alpha_{\mathrm{epi}} = 60°$, which are different from the three possible values of $\tilde{\boldsymbol{\alpha}}$.

In the preceeding work of this thesis, two noise models have been considered: the first was iid Gaussian noise on the 25 standardized outputs (LVV and 24 circumferential strains) and the other considered iid Gaussian noise on the circumferential strains with a unique noise standard deviation of 5 ml on the LVV. Were we to assume this model and sample the material parameters from test data that, like $\boldsymbol{y}^*$, were produced by a process with fibre angles different from those used to generate the training data for the emulator, then the inferred parameters $\hat{\boldsymbol{\theta}}$ would be biased, as discussed in Section 9.1.4. As a result, the following would hold:

$$\tilde{\boldsymbol{y}} \neq \boldsymbol{\eta}(\hat{\boldsymbol{\theta}}, \mathbf{h}). \tag{9.14}$$

where $\tilde{\boldsymbol{y}}$ comes from (9.13). The lack of equality here is solely because $\hat{\boldsymbol{\theta}} \neq \tilde{\boldsymbol{\theta}}$, and ideally

we would infer material parameters $\hat{\boldsymbol{\theta}}$ such that (9.14) is an equality. This is what we hope to achieve by sampling parameters using the likelihood function from (9.6), where the term $\mathbf{b}(\mathbf{x}) + v$ corrects for discrepancy between $\boldsymbol{\eta}(\hat{\boldsymbol{\theta}}, \mathbf{h}) = \tilde{\boldsymbol{y}}$ and $\boldsymbol{y}^*$. Therefore, when sampling parameters using (9.6) as the likelihood function, we can assess the success of the model by comparing the distribution of predictions from the emulator at each posterior sample with $\tilde{\boldsymbol{y}}$. These simulations $\tilde{\boldsymbol{y}}$, which are produced according to (9.13), are now referred to as *benchmark simulations*. Note that we would ideally use the simulator to obtain the distribution of the simulator evaluations over the posterior of the parameters but due to computational costs, the emulator will be used instead.

First, we can perform a posterior predictive check by comparing the posterior predictive distribution with the test data. Since there are many test cases, this is only done for a single example (with $\alpha_{\text{endo}} = 0°, \alpha_{\text{epi}} = 90°$). Figure 9.6 shows the posterior predictive distributions of the circumferential strains in the case where parameters were sampled with the likelihood with model discrepancy (blue dashed lines) and the model that does not account for model discrepancy (red solid lines). Focusing first on the model without discrepancy correction, we see that the variance in the posterior predictive is quite large. Because the discrepancy introduced in $\boldsymbol{y}^*$ is not fully within the capacity of the emulator, the noise standard deviation is overestimated (by a factor of 2) and very little is learnt about the material parameters of the model. For the model with discrepancy correction, the variance in the posterior predictive is smaller. Indeed, the noise variance is now slightly underestimated, with the median of the posterior distribution of the standard deviation equal 0.024, which underestimates the true value of 0.03 as shown in Figure 9.7.
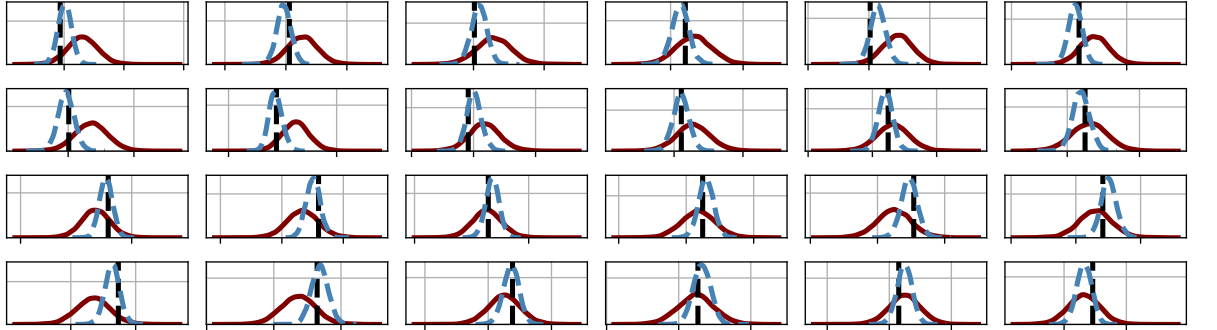

Figure 9.6: Comparing the posterior predictive distribution of the model that does not account for model discrepancy (red solid lines) and the model that does account for model discrepancy (blue dashed). Each subplot gives the result for a different circumferential strain, with the vertical dashed line showing the measured strain value.

We can now assess the performance of the different models by comparing the posterior distribution of $\eta(\boldsymbol{\theta}, \mathbf{h})$ with the benchmark simulations that were discussed previously. This time the volume only model can be considered as well. For all models, a distribution of model predictions is created by predicting from the emulator at each posterior sample.
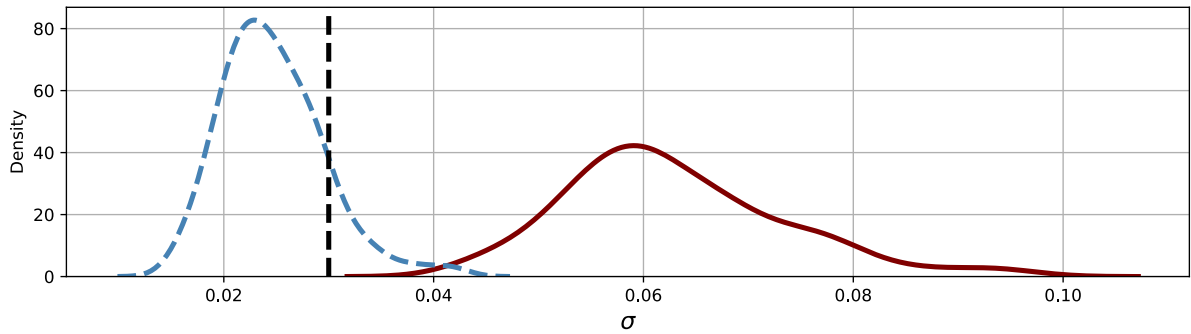
Figure 9.7: Comparing the posterior distribution of the noise standard deviation with the ground truth value, which is indicated by a vertical dashed line at 0.03. Without model discrepancy (red solid line), there is a huge inflation in noise standard deviation because the discrepancy between model and data is outwith the capacity of the model. With the inclusion of model discrepancy, this noise standard deviation is slightly underestimated as the model discrepancy term models some of the extra structure in the data.

According to the discussion at the beginning of this section, this distribution should contain the benchmark simulation if the model discrepancy correction is successful. Figure 9.8 provides detailed results for three different test cases (in each column) using each of the different models (in the different rows), as indicated in the legend. Note that the first column shows results for the same test case for which the posterior predictive distribution was visualized in Figure 9.6. The volume and polar GP models perform similarly, both outperforming the model without any discrepancy correction, which is particularly bad in the last example. The results for all test cases are compared in Figure 9.9, which shows the distribution of errors between the benchmark simulation and the posterior of η over the 12 different test cases. These are all combined into one distribution for each method of inference. From these plots we see that the volume only model and the model with discrepancy correction provided by the polar GP perform almost identically, both of which outperform inference with the standard normal likelihood function ((9.6) without any discrepancy correction included on the circumferential strains).

9.4.4   LV geometry study

An initial look at the effect of LV geometry

It is already clear from previous work that the behaviour of the LV in diastole can vary depending on the LV geometry. This was made evident by the error introduced to the outputs of the simulator when adopting a PCA representation of the LV geometry in Chapter 6. In Figure 9.10, the variations that result from running the simulator at the same material parameters for different LV geometries are presented. From this plot we see that the LV geometry can have a fairly substantial effect on the simulator outputs
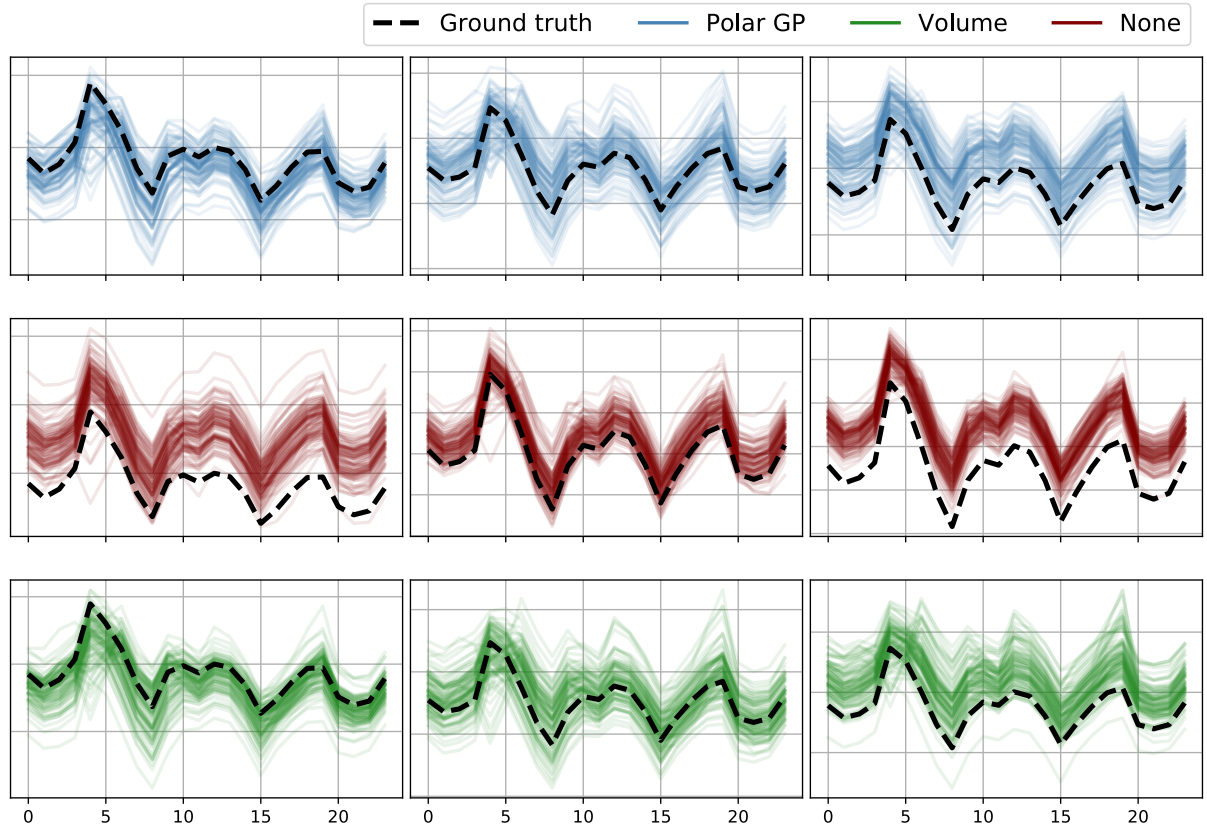
Figure 9.8: Comparing the posterior distribution of $\eta(\boldsymbol{\theta}, \mathbf{h})$ with the benchmark simulations. Three different test cases are shown (in each column), with the different rows giving the results when parameters are inferred using each of the different models, as indicated in the legend. "None" corresponds to the model that includes LVV and circumferential strains without accounting for any model discrepancy.
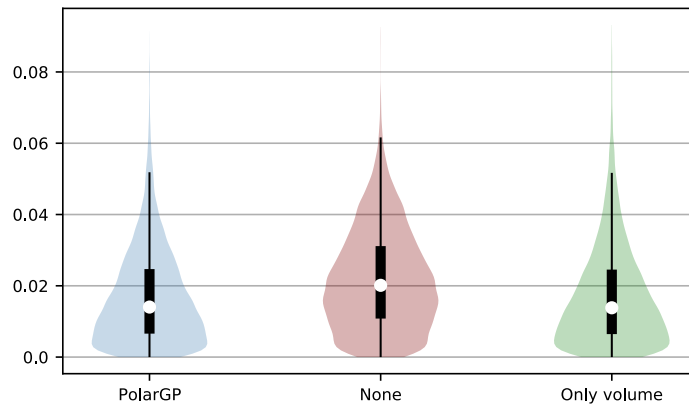

Figure 9.9: Distribution of absolute errors between the posterior distribution of $\eta(\boldsymbol{\theta}, \mathbf{h})$ and the benchmark simulations for all 12 test cases combined, across the three different noise models. The Polar GP and model with only volume perform similarly, both outperforming the standard iid Gaussian noise model (labelled as "None").

and, similar to the fibre angles, the effect of LV geometry variation seems to be most pronounced in the inferior lateral region.
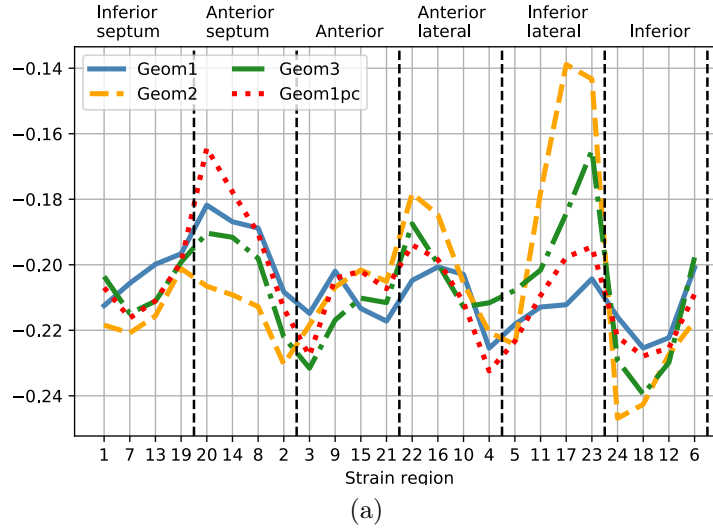

(a)

Figure 9.10: The effect of LV geometry on the simulated circumferential strains and LVV.

Operator effect in the geometry domain

As discussed in Section 9.3.3, 3 operators segmented 6 different LV geometries. A visual comparison of the LV geometry reconstructions of the different operators is provided in Appendix E.1, along with comparisons of the cavity volumes. Quantitative differences between the segmentations are provided in Table 9.4, where each subtable gives the results when a different operator is used as the main operator (recall that the main operator's segmentations are taken to be the benchmark). Also provided is the reconstruction error of a PCA reconstruction of the main operator's LV geometry segmentation, with the values highlighted in bold indicating the number of PCs needed to outperform both of the test operators in a particular example. In general, it appears that the shapes of operator 1 and 3 are in good agreement (at least in terms of MSE), but 8PCs are enough to outperform the test operators in 5 out of 6 test cases. The first segmentation of operator 1 appears to be a bit of an outlier, suggesting some error has been made during the segmentation. Indeed, upon further inspection it was discovered that an obvious error had been made during motion correction (where we align the short and long axis of the LV geometry reconstructions). The bottom row of the table provides the proportion of cases where PCA outperforms either operator. For instance, the value 0.44 below the 5PC column shows that in 44% of the 18 cases, 5PCs give lower reconstruction error than either operator. In all but one example, 10PCs (and often less) is enough to outperform a different operator in terms of reconstruction error.

Table 9.4: Comparing reconstruction errors of three different operators. Each of the 6 rows gives the results for an individual patient segmented by the operators. The main operator's reconstruction is taken to be the benchmark, compared with its PCA reconstruction and those of the other two operators. The first reconstruction with PCA that gives lower MSE than either operator is in bold.

Main operator: 1							
Oper 2	Oper 3	5PC	6PC	7PC	8PC	9PC	10PC
0.025	0.030	**0.002**	0.002	0.002	0.001	0.001	0.001
0.003	0.003	0.012	0.007	0.006	0.005	0.004	**0.003**
0.009	0.006	0.008	0.007	0.007	**0.006**	0.004	0.004
0.020	0.005	**0.005**	0.005	0.004	0.002	0.002	0.002
0.016	0.008	0.013	0.011	**0.007**	0.007	0.006	0.006
0.005	0.005	**0.003**	0.003	0.003	0.003	0.003	0.003
Main operator: 2							
Oper 1	Oper 3	5PC	6PC	7PC	8PC	9PC	10PC
0.025	0.004	**0.003**	0.003	0.002	0.002	0.002	0.002
0.003	0.006	0.014	0.007	0.007	0.006	0.005	0.004
0.009	0.013	**0.003**	0.003	0.003	0.002	0.002	0.002
0.020	0.013	**0.012**	0.012	0.004	0.004	0.004	0.004
0.016	0.023	0.021	0.020	**0.011**	0.010	0.008	0.008
0.005	0.004	**0.002**	0.002	0.002	0.002	0.002	0.002
Main operator: 3							
Oper 1	Oper 2	5PC	6PC	7PC	8PC	9PC	10PC
0.030	0.004	0.006	0.005	**0.002**	0.002	0.002	0.002
0.003	0.006	0.006	0.004	0.004	0.003	**0.002**	0.002
0.006	0.013	0.007	0.007	0.007	**0.005**	0.004	0.004
0.005	0.013	0.010	0.009	0.006	**0.004**	0.003	0.003
0.008	0.023	0.015	0.014	**0.008**	0.008	0.007	0.007
0.005	0.004	**0.002**	0.002	0.002	0.002	0.002	0.002
-	-	0.44	0.44	0.67	0.83	0.89	0.94

Operator effect in the biomechanical domain

The primary interest lies in the effect that the operator induced variations in the LV geometry have on the simulator. First considered using a comparison in output space, the results of this analysis are found in Figure 9.11. As before, each operator is individually treated as a main operator, and the simulated circumferential strains obtained using the main operator's LV geometry are compared with simulated quantities obtained using the 2 test operators and the 5 and 10 PC reconstructions. The labels on the horizontal axis of the violin plots indicate which test reconstruction we are considering. Note that the distributions combine the results for all 6 patients segmented by the operators.

From Figure 9.11, we see that operators 1 and 3 are in quite good agreement: if either 1 or 3 is the main operator and the other is the test operator, the simulated circumferential strain error is roughly the same as that obtained using a 10 PC reconstruction. Operator

2 tends to perform similarly to the 5 PC reconstruction, both of which are worse than the other two test operators. The real assessment of these errors is in the parameter estimation, which is the ultimate interest in this study.
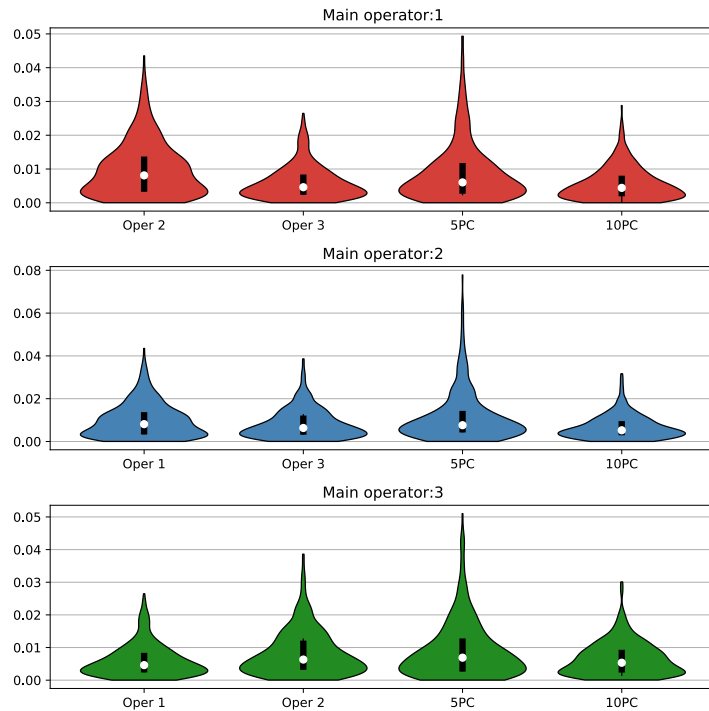

Figure 9.11: Distribution of simulated strain errors with different LV geometry approximations. In each subplot, a different operator is treated as the main operator (as specified in the title) and the simulations with the main operator's LV geometries are compared with the simulations using the other two operator's LV geometries and the 5 and 10 PC reconstructions of the main operator's LV geometry.

The effect of operators on the parameter estimation is explored using the simulation study outlined in Section 9.3.3. In this section the results will be summarized, with the exact numerical results found in Appendix E.2. The plots in Figure 9.12 compare the inference results using the absolute error in each individual parameter, with each row corresponding to a different main operator (recall that the six LV geometry reconstructions obtained by the main operator are used to generated test data for parameter optimization). The results in the plots suggest that optimization using each different test geometry gives similar results no matter the main operator.

A comparison that depends on the Euclidean distance of each individual parameter from its ground truth value is not a fair indicator of performance because it fails to take into account the relationship between the different material parameters. This inter-dependence is indicated by the ridges in Figure 9.13, where the loss function is plotted as a function of different combinations of material parameters. When assessing the success of parameter inference, a perturbation of the parameters by distance D along a ridge should be viewed
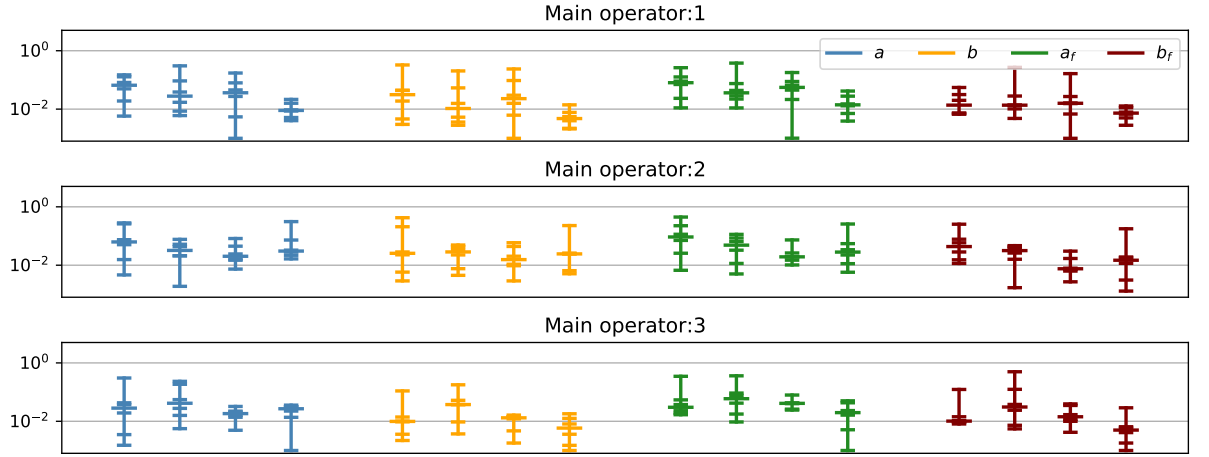
Figure 9.12: Results of parameter estimation. Each subplot shows the absolute errors in material parameters when a particular operator is treated at the main operator, whose segmentation is used to generate synthetic test data. In each subplot, the summaries are split according to each constitutive parameter with the sets of 4 error bars showing results for test operators, 5PC and 10PC. For instance, in the top subplot the first 4 (blue) error bars show the error in parameter estimates for a when using the reconstructions of operator 2, operator 3, 5PC and 10 PC where the segmentations of operator 1 are used to generate the test data. Each error bar shows the results for the 6 different LV geometries as narrow tick marks with the median shown as a wider tick mark.

differently to perturbation by distance D off a ridge. One metric that could overcome this issue is the Mahalanobis distance, which accounts for correlations between variables of interest. However, this is not straightforward to compute because it relies on a numerical estimate of the Hessian to provide an approximate covariance matrix. Instead of this, a metric in simulator output space can be used.

Recall that the data used in this study were generated using the main operator's LV geometry segmentation and the ground truth material parameters $\tilde{\boldsymbol{\theta}}$. For the current comparison, this simulation is referred to as a benchmark simulation. For each test LV geometry, the material parameters were optimized to obtain $\hat{\boldsymbol{\theta}}$, and in general $\hat{\boldsymbol{\theta}} \neq \tilde{\boldsymbol{\theta}}$ because $G \neq \hat{G}$. In the current comparison, these parameters will be compared not in parameter space, but in function space. To do so, the simulator is evaluated using the main operator's LV geometry segmentation and $\hat{\boldsymbol{\theta}}$. This simulation can then be compared with the benchmark simulation to assess the accuracy of inference. By assessing the accuracy of the parameter estimates in simulator output space, movements along a ridge in parameter space will be penalized less than movements off a ridge.

The results of the output comparison are shown in Figure 9.14. Each subplot shows the results when a different operator is treated as the main operator, as indicated by the titles of the plots. The first two distributions correspond to results with the 2 test operators (indicated by the label of the violin plot) and the final 2 distributions are those
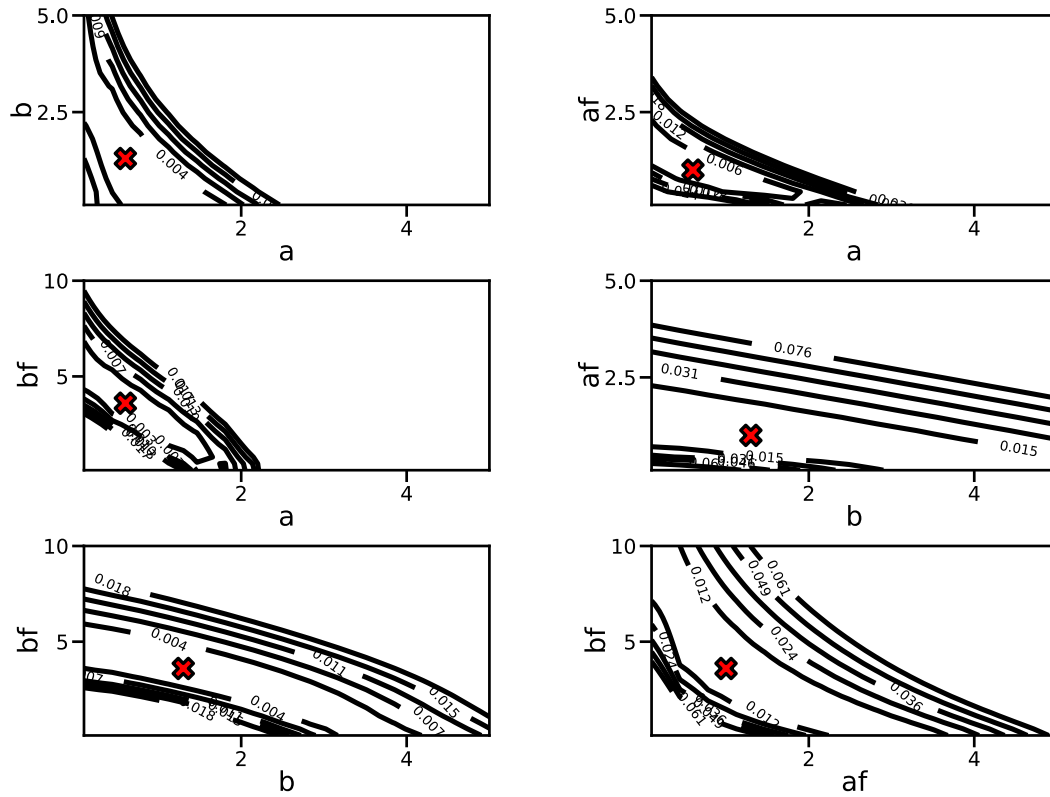
Figure 9.13: Contour plot of the RSS function over different combinations of parameters. The red X marks the ground truth parameter value.

corresponding to 5 and 10PC reconstructions of the LV geometry. Although it is difficult to see the details of some of the distributions, it is clear that the two different PC representations provide results that are at least as consistent as the new operators. This result suggests that we will not gain anything by including large numbers of PCs in the parameter inference scheme, because the errors introduced by new operators lead to larger errors than with just 5 PCs.

9.5   Discussion and conclusion

Having acknowledged an obvious mismatch between the cardio-mechanic model and measured CMR data in Section 9.1.1, the work in this chapter considered the effect of erroneous physiological conditions on the model calibration task. The first study showed the change in the fit of the model to the measured data when the fibre angles are allowed to vary during optimization. Building on this, a synthetic data study was carried out where test data were generated using deliberately incorrect angles in the rule-based fibre generation
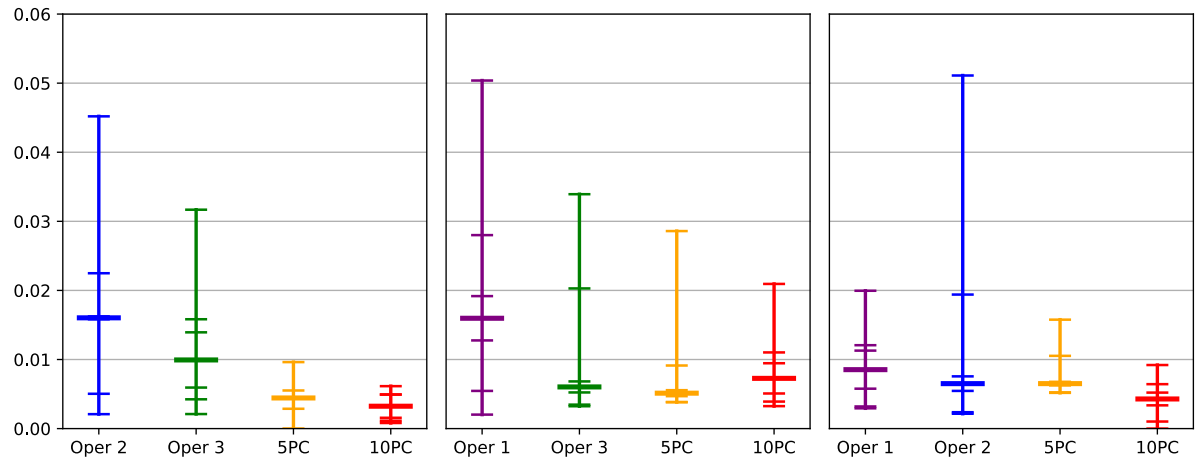
Figure 9.14: Comparing the parameter estimates in output space using Euclidean distance between circumferential strains. Each plot gives the results for a different main operator. The simulator is evaluated, with the ground truth LV geometry, at the optimized parameters (the test simulation) and ground truth parameters (the benchmark simulation). Comparing the test and benchmark simulation allows for a comparison of parameter estimates that takes into account correlations between the parameters.

algorithm. Finally, the effect of the LV geometry on the parameter estimation task was considered. In particular, the knock-on effect of LV segmentation operator errors on the parameter estimation was studied.

In Section 9.4.2, the fibre angles were treated as extra variables during inference and jointly optimized along with the material parameters. This was motivated by the fact that the volume tends to be underestimated during inference, and the exploratory analysis in Figure 9.5 that showed that a change to the fibre angles can affect the circumferential strains more than the cavity volume. Therefore, it was postulated that one reason for the LVV-circumferential strain mismatch could be the fibre angles. This joint-optimization study showed that tuning the parameters and the angles together does allow for an improvement in the fit to the measured CMR data. These results should motivate future work on inferring the material parameters and fibre angles jointly, especially given the difficult of measuring the fibre angles in vivo. To do so reliably would require more precise or more plentiful datasets than those that are currently available.

The next stage in studying the effect of fibre angles was to consider a synthetic data study. In replacement of the cardio-mechanical model, an emulator was used where the training data had been generated assuming a fixed set of fibre angles (which was the case with all other emulators used for I-UQ in this thesis). To introduce model discrepancy, the synthetic data were generated using fibre angles different from the ones used when generating the training set. This introduced a mismatch between the emulated circumferential strains and the measured circumferential strains, but resulted in no systematic mismatch in the LVV. In total, three different models were used to estimate the parameters from

the test data. The first was the standard Gaussian noise model used in the previous work in this thesis. In particular, this assumed a fixed noise variance on the LVV and iid Gaussian noise on the circumferential strains, the latter of which could be inferred during inference. The second was a model that accounted for discrepancy on the circumferential strains via a correction function on the circumferential strains that was modelled using a GP as outlined in Section 9.2.2. To ensure model identifiability, the LVV was assumed to be corrupted with Gaussian noise with a fixed noise variance, but no model discrepancy. Finally, a model that included the LVV alone was considered, again with fixed noise standard deviation. The fixed noise standard deviation for the LVV was the same across the three models. By assuming the first noise model, a bias was introduced to the parameter estimates as a result of the mismatch between the fibre angles used in the emulator and the fibre angles used to generate the test data. In addition, there was an inflation in the inferred noise variance, resulting from the fact that there was extra discrepancy between model and data that could not be accounted for by the model. This discrepancy was corrected by including a discrepancy term on the circumferential strains, allowing for a reduction in the inferred noise variance on the circumferential strains that was closer to the ground truth value, and a reduction in the bias of the parameter estimates. Lastly, the model with discrepancy correction and the model with only LVV were shown to provide similar inference results. While this might motivate the idea of using the LVV alone for parameter inference, the benefit of the model with discrepancy correction is that it is automatic, without requiring us to study the relationship between circumferential strains and LVV. In the future, it would be worth exploring ways of incorporating a more informative disrepancy correction for both the circumferential strains and the LVV.

One of the main tasks in this thesis has been to allow for the generalization of a statistical emulator to different LV geometries. Previous work, in particular Chapter 6, suggested that 5 PCs were not sufficient for representation of the LV geometry in parameter inference. Making use of three different LV segmentation operators, the final study in this chapter has compared inter-operator LV geometry errors with those obtained using PCA. The comparison was performed first in geometry space and then using the biomechanical model. Considering first the geometry space comparison, the results in Table 9.4 showed that in all but one example, a 10 PC reconstruction gave a more accurate approximation to an LV geometry than a segmentation obtained by a new operator. The next stage of the comparison considered the biomechanical model, first in terms of the forward model and then in terms of inverse modelling. In the forward model, operators 1 and 3 were found to be fairly consistent, both achieving results similar to 10 PCs. On the other hand, Operator 2 and the 5 PC reconstructions were noticably worse. The real test of these different LV geometry approximation methods was in the parameter estimation, which is ultimately our main interest. Comparing parameter optimization results in parameter

space, the different LV geometry approximations were found to give similar results. It was suggested that this parameter space comparison might be suboptimal because it does not account for correlations between parameters. As a result, an output space comparison of the parameter estimates was performed. The result of this comparison showed that 5 and 10 PCs give better results than the new operators, which is a quite substantial result in the context of this thesis. In particular, this suggests that there is no advantage to improving the accuracy of the LV geometry reconstruction in the generalized emulator.

While the results in the current chapter suggest that there is no need to improve the accuracy of the generalized emulators that were built in this thesis, it is possible that the accuracy of LV segmentation will improve to be more objective in the future. In this case, a more accurate emulator would be beneficial for parameter inference. Appendix F suggests an alternative approach to emulation of the simulator model, where a set of basis functions is learnt from a training set of LV geometries (with corresponding simulations) and adapted to a new test subject. Fitting to the new test subject relies on a small number of simulations that can be run in parallel, but preliminary results show that the model can be very accurate with only 20 simulations. Further exploration of the design and training of the model would likely lead to further improvements in accuracy.

In conclusion, the present study has highlighted several problems with the translation of mathematical models to the clinic. First of all, the large mismatch between measured data and the LV model was highlighted, acknowledging that this could be the result of either model-reality descrepancy or noise introduced by the measurement process. The former situation was studied in the remainder of the chapter, which sheds light on the effect that the fibre angles of the rule based method and the error introduced by operators to the LV geometry can have on the parameter estimation. In particular, a model that accounts for model discrepancy was found to provide better results than one that assumes iid measurement errors, while providing no noticeable difference with a model that included LVV alone. This study assumed that there is no bias introduced to the LVV, which was the case for the fibre angles used to generate the test data. In the LV geometry study, the effect that operator segmentation variation has on the parameter estimation in the LV model was found to exceed that that was introduced by PCA. This should motivate mutliple lines of future study. Principal among these is the collection of larger and more reliable datasets for the estimation of the model parameters, without which the parameter estimates obtained for these models cannot really be trusted. Secondly, a better informed discrepancy correction that accounts for bias in the LVV and circumferential strains would allow for more trustworthy inferences than are currently available.