Research

Cite this article: Del Corso G, Verzicco R, Viola F. 2020 Sensitivity analysis of an electrophysiology model for the left ventricle. *J. R. Soc. Interface* **17**: 20200532. <http://dx.doi.org/10.1098/rsif.2020.0532>

Received: 7 July 2020

Accepted: 2 October 2020

Subject Category:

Life Sciences—Mathematics interface

Subject Areas:

biomathematics, biomechanics

Keywords:

uncertainty quantification, polynomial chaos expansion, electrophysiology, global sensitivity analysis, bidomain model

Author for correspondence:

Francesco Viola

e-mail: francesco.viola@gssi.it

Sensitivity analysis of an electrophysiology model for the left ventricle

Giulio Del Corso¹, Roberto Verzicco^{1,2,3} and Francesco Viola¹

¹Gran Sasso Science Institute (GSSI), L'Aquila, Italy

²University of Rome Tor Vergata, Rome, Italy

³POF Group, University of Twente, Enschede, The Netherlands

GDC, 0000-0003-4604-2006; RV, 0000-0002-2690-9998; FV, 0000-0003-1303-5934

Modelling the cardiac electrophysiology entails dealing with the uncertainties related to the input parameters such as the heart geometry and the electrical conductivities of the tissues, thus calling for an uncertainty quantification (UQ) of the results. Since the chambers of the heart have different shapes and tissues, in order to make the problem affordable, here we focus on the left ventricle with the aim of identifying which of the uncertain inputs mostly affect its electrophysiology. In a first phase, the uncertainty of the input parameters is evaluated using data available from the literature and the output quantities of interest (QoIs) of the problem are defined. According to the polynomial chaos expansion, a training dataset is then created by sampling the parameter space using a quasi-Monte Carlo method whereas a smaller independent dataset is used for the validation of the resulting metamodel. The latter is exploited to run a global sensitivity analysis with nonlinear variance-based indices and thus reduce the input parameter space accordingly. Thereafter, the uncertainty probability distribution of the QoIs are evaluated using a direct UQ strategy on a larger dataset and the results discussed in the light of the medical knowledge.

1. Introduction

Advances in computational science have enabled the development of digital twins of biological systems of which a popular application is in cardiology and, more in general, in cardiovascular medicine. This approach not only improves the predicting capabilities of diagnostic tools and gains more insight into patients' pathologies, but also provides a numerical framework to test innovative medical devices and to refine them before running further *in vivo* experiments on animals or humans. Predictive mathematical models imply a physical basis and the quantities of interest (QoIs) have to be obtained by solving the governing equations for the system under investigation.

In particular, modelling human heart functioning entails reproducing the complex electrical activation triggering muscular contraction, which includes a conductive network that propagates the local transmembrane potential of the myocytes from the sinoatrial node (SA-node) to the atrial and the ventricular muscles. As visible in figure 1a, depolarization of the local myocytes originates in the SA-node that is located in the right atrium close to the entrance of the superior vena cava [2], and then propagates across the right atrium and reaches the atrioventricular node (AV-node). After the transmembrane depolarization front leaves the AV-node, it spreads towards the ventricles passing through the His bundle that bifurcates into a set of bundles that conducts the signal to a fast conduction system (the Purkinje network), thus allowing for more uniform signal propagation in the ventricular myocardium and for subsequent simultaneous muscular activation and efficient blood pumping towards the circulatory system. As the transmembrane depolarization front reaches a certain location in the myocardium the local transmembrane potential of the myocytes rapidly changes from the negative potential (of about -85 mV) to a positive value (of about 20 mV) before returning to the resting negative

Figure 1. (a) Sketch of the electrical network of the heart adapted from [1], with highlighted left ventricle components. (b) Typical depolarization/polarization cycle of a cardiac myocytes (action potential and intracellular calcium profiles), which triggers the muscular active tension. (c) ECG pattern in a healthy subject, the ventricular depolarization (QRS complex) and repolarization (T wave) patterns are indicated by the red line.

potential after about 300 ms, as indicated in figure 1b. This transient depolarization of the cardiomyocytes (*action potential*) yields the release of cytosolic calcium from the sarcoplasmic reticulum that originate a contractile force within the cardiac muscle cells (the sarcomers), thus causing the myocardium to contract. The contraction of the muscular fibres can be detected placing electrodes on the patient's limbs and chest as done in electrocardiography (ECG), where the P wave represents the depolarization of the atria, the QRS complex corresponds to the depolarization of the ventricles and the T wave indicates repolarization of the ventricles; see figure 1c for a typical healthy ECG patterns. Variations of the normal ECG pattern and durations are usually ascribed to cardiac pathologies and abnormalities, such as atrial fibrillation and ventricular tachycardia.

This electrical activation of the heart has been studied using different approaches depending on the application. In the interconnected cable method [3,4], the cardiac tissue is modelled as a connected network of discrete cables representing muscle cardiac fibres. This approach is broadly used because it accurately describes the fibres electrical activation at a lower computational cost compared to other methods and can be efficiently parallelized [5]. Starting from the cable equation, a fractional Laplacian formulation can be used to model the cardiac tissue including the macroscopic effects of structural heterogeneity on impulse propagation [6] or to incorporate more complex conduction structures, such as cardiomyocytic fibre orientation and the His–Purkinje activation network [7]. The propagation of the cellular depolarization front can also be modelled by an eikonal approach, where the excitation time, defined at each point mesh as the time instant at which the transmembrane potential crosses the value midway between its resting and plateau potentials, is solved [8]. On the other hand, the bidomain model, so called because the conductive myocardium is modelled as an intracellular and an extracellular overlapping continuum media separated by the myocytes membranes, computes the electrical cardiac activity by solving the quasi-steady electric equations [9,10]. The resulting system of reaction–diffusion

partial differential equations governs the electrical propagation across the myocytes and is coupled with a set of ordinary differential equations (ODEs; the cellular ionic model) describing the current flow through the ion channels, see § 2. In the case that the extracellular and intracellular conductivity tensors are parallel to each other, the bidomain equations can be simplified as a single governing equation for the transmembrane potential, the monodomain system, which is computationally cheaper not only because the number of degrees of freedom is reduced but also because the equations are more stable numerically [11]. Differences in the electrical propagation and intra/extracellular potentials between monodomain and bidomain are typically very small unless complex activation patterns are applied, as in the case of pacing or defibrillation, where a monodomain approach is generally deprecated [12]. The bidomain model is hence the state-of-the-art mathematical model for reproducing the electrical activation of the heart chambers in healthy and pathological conditions [11,13] and the propagation of the myocytes depolarization over the cardiac tissue (see [13,14] among others). Furthermore, this model is seen to reproduce cardiac phenomena including ischaemic events and defibrillation [15] and it has been validated through animal experiments [16,17]. More recently, the bidomain model has been coupled with a fluid–structure solver to build a multiphysics model for the left heart including the mitral and aortic valves with the aim of using computer simulations to evaluate medical quantities that would be exceedingly difficult or impossible to be measured *in vivo* or *in vitro* [18].

Although the bidomain model is nowadays a well-established tool to study cardiac electrophysiology, several quantities have to be provided as input parameters such as the chambers' geometry, the orientation of the muscular fibres, the mechanical properties of the biological tissues (the behaviour of which is nonlinear and orthotropic) along with the conduction velocities within the electrical network, just to cite a few. Only some of these quantities can be estimated *in vivo* through the scanning methods (e.g. echocardiography, MRI and CT) and a significant variability among individuals is known to

exist [19]. Calibrating the computational models is crucial for the personalization of the therapies and, as an example, for the prediction acute haemodynamic changes associated with cardiac resynchronization therapy (CRT) [20]. However, the sparsity and the noise of the clinical data used to calibrate the model parameters have a major impact on the results of the model including the transmembrane potential propagation in the cardiac tissue [21]. In this framework, several computational techniques for model calibration have been introduced, such as variational approaches based on the constraint optimization [22,23], data assimilation methods [24] or even patient-specific Bayesian inference strategies using polynomial chaos expansion [25]. The uncertainty regarding the input parameters opens the question about the reliability of the model results, thus calling for a rigorous uncertainty quantification (UQ) analysis. The latter provides a set of mathematical methods to study the uncertainty propagation of the input parameters of an electrophysiology model for the heart on the model results, by combining the deterministic approach used to solve the PDEs of the physical model with a probabilistic framework to handle the uncertainties of input parameters and QoIs.

In this work, we apply UQ analysis to an electrophysiology computational model based on the bidomain equations. However, considering whole-heart electrophysiology would require that the uncertainties of the fast conductivity bundles (e.g. internodal pathways, His bundle, left–right posterior–anterior bundles), of the complex geometry of the heart chambers and of the heterogeneous properties of the cellular model through the myocardium are dealt with simultaneously. Hence, the high number of the uncertain parameters along with the computational cost of a single electrophysiology simulation of the full heart would make the UQ analysis very demanding and it should be tackled by successive steps. With this goal in mind, we have decided to focus here on the UQ analysis of the electrophysiology of the left ventricle (LV), which is of paramount importance since the LV is the heart chamber with the thicker muscular myocardium that pumps oxygenated blood to the global body circulation and, as a consequence, becomes diseased more frequently with important practical and clinical implications. Furthermore, simulating LV electrophysiology allows the consideration of the relevant electrophysiology features such as the fibre orientations, the velocity of the electrical conduction over the myocardium, the variability in chamber geometry and the properties of the cellular model. Although we mainly consider here QoIs relative to wavefront propagation in a healthy ventricle that could have been investigated with a simplified and computationally cheaper propagation model (e.g. eikonal or cable models), a bidomain approach has been adopted so that the present analysis can be extended to the case of cardiac pacing, defibrillation and arrhythmia using the same UQ methodology.

In the preliminary phase of this UQ analysis, the probability distribution functions (PDFs) corresponding to the uncertainty of these input parameters are determined. However, the main difficulty arises here as there are no data available from the literature to estimate their PDFs of the cellular model input. For this reason, the UQ analysis has been divided in two parts by studying separately the effect on LV electrical activation (i) of the geometrical chamber parameters and of the electrical conductivities, whose uncertainty could be estimated by experimental measures reported in the literature, and (ii) of the cellular model inputs, whose PDFs are unknown.

Specifically, in §3, a global sensitivity analysis [26] of the effects of the ventricular geometry and of the electrical conductivities on the myocardium depolarization is carried out. In this analysis, the uncertain geometrical and electrical input parameters of the model are identified and their PDFs are estimated from available data from the literature. The sensitivity analysis is carried out using both a direct approach (QMC sampling to compute Sobol' indices using Saltelli's algorithm) and a metamodel one (adaptive polynomial chaos expansion, PCE), which has also been used to run a forward sensitivity analysis to produce the PDFs of the QoIs. Owing to the great computational advantage of using a smaller dataset than direct strategies, the PCE is becoming a common tool in electrophysiological UQ [27,28] and can be applied to estimate the effect of uncertainty on different QoIs (such as steady-state activation/inactivation and current density in ionic channel models [28]), or to speed up the computation eikonal models using a Bayesian multifidelity approach by integrating the electrophysiology solver with a metamodel to achieve near real-time UQ analyses [29].

The second analysis reported in §4, aims to understand the role of the input parameters of the ten Tusscher–Panfilov cellular model [30] governing the electrical current in the ionic channels of the myocytes on the action potential profile as well as on the ventricular electrical activation. The effect of the input electrical conductivities and ions concentration is rationalized through a local sensitivity analysis (as defined in [31,32]) using a metamodel technique based on an adaptive sparse PCE. This analysis identifies the most relevant input quantities of the cellular model, thus allowing for a significative model reduction of the uncertain input parameter space and suggesting a control strategy to adapt the cellular model to pathological cases. Conclusions and perspectives for future works are then given in §5.

2. Problem configuration and numerical method

The electrophysiology model and the numerical method used in this study are thoroughly described in [18] and, therefore, only the main features are summarized here. The bidomain equations governing the depolarization of the ventricular myocardium read

$$\begin{aligned} \chi \left(C_m \frac{\partial v}{\partial t} + I_{ion}(v, s) + I_s \right) &= \nabla \cdot (\mathcal{M}_{int} \nabla v) + \nabla \cdot (\mathcal{M}_{int} \nabla v_{ext}), \\ 0 &= \nabla \cdot (\mathcal{M}_{int} \nabla v + (\mathcal{M}_{int} + \mathcal{M}_{ext}) \nabla v_{ext}), \\ \frac{\partial s}{\partial t} &= F(v, s), \end{aligned} \quad (2.1)$$

where v and v_{ext} are the unknown transmembrane and extracellular potential (expressed in mV) while the surface-to-volume ratio of cells, $\chi = 1400 \text{ cm}^{-1}$ along with the membrane capacitance $C_m = 1 \text{ } \mu\text{F cm}^{-2}$ are set as in [33]. The symbols $\nabla \bullet$ and $\nabla \cdot \bullet$ indicate the gradient and divergence operators, respectively, whereas $\partial \bullet / \partial t$ indicates the partial derivative with respect to time. $\mathcal{M}_{int}$ and $\mathcal{M}_{ext}$ are the conductivity tensors of the intracellular and extracellular media that depend on the local fibre orientation with a faster propagation velocity along the fibre direction than in the normal ones. These tensors are diagonal when expressed in the local coordinates (i.e. fibre, sheet and sheet-normal

Figure 2. Snapshots of the transmembrane potential as a function of time showing the left ventricle depolarization. Top row: perspective side view. Bottom row: view from the ventricle apex.

directions)

$$\mathcal{M}_{\text{ext}}^* = \begin{bmatrix} M_{\text{ext}}^{\parallel} & 0 & 0 \\ 0 & M_{\text{ext}}^{\perp} & 0 \\ 0 & 0 & M_{\text{ext}}^{\times} \end{bmatrix} \quad \text{and} \quad \mathcal{M}_{\text{int}}^* = \begin{bmatrix} M_{\text{int}}^{\parallel} & 0 & 0 \\ 0 & M_{\text{int}}^{\perp} & 0 \\ 0 & 0 & M_{\text{int}}^{\times} \end{bmatrix}, \quad (2.2)$$

and the corresponding global conductivity tensors are given by $\mathcal{M}_{\text{ext}} = \mathcal{A} \mathcal{M}_{\text{ext}}^* \mathcal{A}^T$ and $\mathcal{M}_{\text{int}} = \mathcal{A} \mathcal{M}_{\text{int}}^* \mathcal{A}^T$, where $\mathcal{A}$ is the rotation matrix containing column-wise the components of fibre, sheet and sheet-normal unit vectors and $\mathcal{A}^T$ is its transpose [11]. The last of equations (2.1) indicates a cellular model governing the transmembrane ionic currents along with ionic concentrations and ion channel kinetics. The state vector s couples the cellular model with the bidomain equations through the ionic current per unit cell membrane I_{ion} (measured in $\mu\text{A cm}^{-2}$). Among the several models that exist to represent the cellular-scale dynamics in the bidomain equations, the ten Tusscher–Panfilov cellular model is used, which is seen to correctly reproduce the action potential within ventricular myocytes with physiological detail [30]. The model includes all the major ion channels as well as intracellular calcium dynamics and it consists of a nonlinear system of 19 ODEs, which are not reported here for the sake of brevity and that are indicated in compact form in the last equation of the system (2.1). Hence, the ionic current, I_s corresponds to the electrical stimulus applied at the bundle of His initiating the electrical propagation and triggering the ventricular depolarization.

The set of equations (2.1) are discretized on a triangular mesh with Lagrangian finite elements using the electrophysiology library cbcbeat [34], based on the FEniCS FEM library [35,36], which provides an efficient framework to solve electrophysiology models over arbitrary computational domains (see also [18] where further verification and validation tests are reported). The bidomain equations are marched in time through a second-order Strang splitting method where a step accounting only for the ionic and the external currents and another one involving only the right-hand side of equation (2.2) are solved sequentially [11]. A Rush–Larsen scheme for

the ODEs of the cellular model is combined with a $\theta = 1/2$ second-order Crank–Nicholson scheme for the integration of the PDE step. The resulting average CPU time cost to solve a complete depolarization/repolarization cycle of the ventricle on a grid of 4311 cells (corresponding to 45 885 degrees of freedom including those of the cellular model) and using a time step of $dt = 10^{-2}$ ms is 50 CPU-minutes (defined as the time it takes to run the program on a 1 GHz reference processor). The computational resources used for the analysis comprise an Intel Xeon Processor with 16 cores (E5-2620 v3—15 M Cache, 2.40 GHz) that run the same number of simulations simultaneously. Figure 2 shows the electrical activation of the LV at several instants within a heart beat. Initially, all the cells are relaxed with a negative transmembrane potential of about -85 mV and, at a given initial time, an electrical impulse originated at the His bundle propagates in the ventricular muscle causing the cells to locally change the transmembrane potential. As a consequence, a rapid flux of positive ions through the cell membrane occurs, and the transmembrane potential increases to positive values, as visible from the isocontours of the transmembrane potential reported in figure 2*a*. This local depolarization results in a propagating wavefront that travels across the myocardium (see figure 2*b* and *c* activating the whole ventricle as visible in figure 2*d*). As a remark, the myocardium is simplified by considering it as a uniform two-dimensional conductive medium and the bidomain equations are formulated as surface PDEs with anisotropy present only in the tangent plane. Although this approach ignores the transmural anisotropy and corresponds to an overestimation of the transmural speed of depolarization, the surface bidomain model is seen to provide the correct depolarization timings of the ventricles as also observed for the atria in previous works [37,38]. The lack of the fast conductivity fibres has been accounted by scaling the electrical conductivities by a fixed parameter (equal to 4.99, which has been obtained through an optimization procedure based on Brent’s method) so that the computational model reproduces the benchmark timings of ventricular depolarization [1]. The resulting time evolution of the transmembrane potential for one representative point was reported in figure 1*b* together with the intracellular calcium profile triggering muscular contraction. In particular, when the cell is excited by an electrical stimulus over a

Figure 3. (a) Graphical representation of long and short axes of the left ventricle. (b) Orientation of local fibres in the ventricular myocardium, the inset indicates the conduction directions parallel and perpendicular to the fibres.

Figure 4. (a) Side and (b) bottom view of the ventricular activation map $t_a(\mathbf{x})$, corresponding to the first activation time t at any location $\mathbf{x}$. The yellow dot indicates the stimulus position corresponding to the location of the bundle of His.

threshold potential (of about -70 mV), ionic channels at the membrane open and close in a coordinated manner causing the transmembrane potential to rise from its resting negative value to positive values of about 40 mV. This cell depolarization phenomenon is very fast (≈ 2 ms) and is followed by a spike (phase 1) and a relatively long plateau (or dome, phase 2) lasting about 200 – 300 ms, which ends with repolarization (phase 3) to the rest potential (phase 4).

The electrical stimulus initiating the ventricular depolarization is given by $I_s = \tilde{I}_s e^{-\|\mathbf{x} - \mathbf{x}_0\|^2 / \sigma_s^2}$ where the current amplitude is equal to $\tilde{I}_s = 25 \mu\text{A cm}^{-2}$ if the simulation time within a heart beat is less than the temporal duration of the stimulus (5 ms) and null otherwise. The parameter $\sigma_s^2 = 0.5 \text{ cm}^2$ concentrates the stimulus around the position $\mathbf{x}_0$ so that the stimulus is only active in the spatial locations $\mathbf{x}$ close to it. The centre of the electrical stimulus, $\mathbf{x}_0$ is thus placed at the bundle of His position, see figure 4a, where the electrical impulses from the atria are transmitted to the ventricles and which also corresponds to the location where the pacemaker are often implanted (the so-called His-bundle pacing [39]).

3. UQ analysis 1: sensitivity of the electrical activation of the left ventricle on the chamber geometry and electrical conductivities

In this analysis, we investigate the effect of LV geometry and of the myocardial electrical conductivities on ventricular depolarization.

3.1. Input parameters and their PDFs calibrations

The shape of the heart chambers as well as the electrical conductivities are known to vary among individuals and we consider here the effect of the following six input parameters on the ventricular depolarization (figure 3): the long axis of the ventricle L , the sphericity index SI defined as the short to long axis ratio, the intracellular $\mathcal{M}_{\text{int}}$ (parallel and perpendicular) and extracellular $\mathcal{M}_{\text{ext}}$ (parallel and perpendicular) conductivity tensor that depend on local fibre orientation according to equation (2.2). However, before investigating the sensitivity and the uncertainty of the model's results, the uncertainty PDFs of the input parameters have to be

Table 1. Input Gaussian distributions of the model parameter given as mean and standard deviation. For each parameter, the truncation bounds to avoid unrealistic shapes and conductivities are also reported, whereas the truncation probability (i.e. the probability of the truncated tails) is reported in the last column.

input parameters	normal $\mu \pm \text{s.d.}$	truncation bounds	truncation probability
long axis (mm)	80 ± 9	$[\mu/1.3 = 61.53, \mu \times 1.3 = 104.00]$	0.024
SI (sphericity index)	0.52 ± 0.06	$[\mu/1.3 = 0.400, \mu \times 1.3 = 0.676]$	0.027
$M_{\text{int}}^{\parallel}$ (mS mm ⁻¹)	0.268 ± 0.081	$[\mu/5 = 0.134, \mu \times 5 = 0.536]$	0.0496
$M_{\text{int}}^{\perp}$ (mS mm ⁻¹)	0.031 ± 0.0168	$[\mu/5 = 0.0062, \mu \times 5 = 0.155]$	0.070
$M_{\text{ext}}^{\parallel}$ (mS mm ⁻¹)	0.292 ± 0.194	$[\mu/5 = 0.058, \mu \times 5 = 1.46]$	0.1145
$M_{\text{ext}}^{\perp}$ (mS mm ⁻¹)	0.141 ± 0.0687	$[\mu/5 = 0.028, \mu \times 5 = 0.705]$	0.050

Table 2. Available data from the literature on human electrical conductivities.

conductivity (mS mm ⁻¹)	Clerc [43]	Robert <i>et al.</i> [44]	Robert & Scher [45]	Roth [46]	Roth [47]
$M_{\text{int}}^{\parallel}$	0.17	0.28	0.34	0.35	0.2
$M_{\text{int}}^{\perp}, M_{\text{int}}^{\times}$	0.019	0.026	0.06	0.03	0.02
$M_{\text{ext}}^{\parallel}$	0.62	0.22	0.12	0.30	0.2
$M_{\text{ext}}^{\perp}, M_{\text{ext}}^{\times}$	0.24	0.13	0.08	0.18	0.08

known. In [40], the ventricular end diastolic long and short axes (defined as the distance from the apex to the mid-point of the mitral valve and as the length of the segment that perpendicularly intersects the mid-point of the long axis) have been measured using transthoracic endocardiography for a group of 26 men and 26 women with a mean age of 43 ± 14 years for a total of 52 healthy subjects. The resulting distributions of L and SI were observed to follow the normal distributions reported in table 1, which are used here to define the uncertainty PDF of these input parameters of the electrophysiology model.

On the other hand, fewer data on the electrical properties of human myocardium exist owing to the difficulty to measure these quantities *in vivo* and they are insufficient to accurately estimate a PDF for the uncertain conductivities ($M_{\text{int}}^{\parallel}, M_{\text{int}}^{\perp}, M_{\text{ext}}^{\parallel}, M_{\text{ext}}^{\perp}$). A possible UQ strategy that is adopted when data on the input parameters are missing, consists of estimating an initial PDF using the few data available and then refine it through a Bayesian inverse calibration [41]. Also this method, however, relies on an iterative minimization method based on well-known experimental observables, which are usually lacking in the electrophysiology of the human heart. Nevertheless, more biological data on the intracellular and extracellular conductivities have been acquired in the case of animal myocardium in terms of the first two statistical moments, i.e. mean and variance [42]. Although the mean values and the standard deviations of the PDFs can vary among different mammals, we assume human electrical conductivities to be Gaussian distributed and use the few data available from the literature (reported in table 2) to fit these PDFs. In particular, according to the Kolmogorov–Smirnov test [48] the uncertainty of the intracellular and extracellular conductivities corresponds to the normal distributions reported in table 1 (all the p -values are greater than 0.5). Furthermore, in order to avoid unphysical ventricular shapes and

conductivity values, the PDFs of the input parameters have been truncated (table 1) according to the following method. Given a random variable X with PDF $f(X)$ and CDF (cumulative distribution function) $F(X)$, the PDF support is reduced to the range $[a, b]$ by calculating the conditional distribution $g(x) := f(x)|a \leq X \leq b = \mathbb{1}_{[a,b]}f(X)/(F(b) - F(a))$ where $\mathbb{1}_{[a,b]}$ is the indicator function and $g(X)$ is the truncated PDF. In the case, $f(x)$ is the PDF a Gaussian distribution $\mathcal{N}(\mu, \sigma)$, the corresponding truncated distribution is given by the analytical formula $g(x, \mu, \sigma, a, b) = I_{[a,b]} (1/\sigma)(\phi((x-\mu)/\sigma))/(\Phi((b-\mu)/\sigma) - \Phi((a-\mu)/\sigma))$, where ϕ and Φ are the PDF and CDF of the normal distribution $\mathcal{N}(0, 1)$. Adding a truncation bound to the input distributions is important in order to avoid unphysical configurations in the dataset, which could yield numerical instabilities while its effect on the UQ analysis is expected to be negligible as the Sobol' indices are integral quantities that are not sensitive to few tail elements.

Although other types of distributions with bounded support could be considered to fit the discrete distribution of the experimental data available, such as the von Mises or the beta distributions, we have used a truncated Gaussian because it corresponds to the least informative setting as it maximizes the entropy for a fixed mean and variance, with the random variate constrained to be in the interval [49]. Additionally, truncated Gaussians could also account for asymmetries in the PDF shape.

3.2. QoIs

At this stage, the QoIs of the problem are identified by translating clinically relevant quantities measured *in vivo* into outputs of the computational model. The QoIs investigated in this analysis are the *depolarization time* of the LV (DT hereafter, that is expected to be correlated to the conduction velocity) and the *depolarization uniformity* (DU), which are relevant for the patient's health because they are known to be

Figure 5. Scatter plots of the ventricular depolarization (a) time and (b) uniformity against the six input variables computed using the 2500 independent samples of the 20 000 Saltelli's dataset.

associated with the efficiency of the ventricular pumping during systole [1]. Indeed, only a physiological DT and DU determine a timely and almost simultaneous contraction of the muscular fibres of the myocardium, which is needed to effectively propel oxygenated blood from the LV towards the aorta during systole [18]. Given the activation map $t_a(\mathbf{x})$ of the ventricle, which stores the first activation time at each point $\mathbf{x}$ (computed when the transmembrane potential overcomes a threshold of 0 mV, see figure 4), the DT of the ventricular myocardium is defined as

$$\text{depolarization time} = \max\{t_a(\mathbf{x})\}, \quad (3.1)$$

and the DU reads

$$\text{depolarization uniformity} = \text{std}\{t_a(\mathbf{x})\}, \quad (3.2)$$

with both quantities measured in milliseconds and the operators max and std indicating the maximum and the standard deviation of their argument. Note that according to equation (3.2) a higher (lower) value of the second QoI corresponds to a less (more) uniform and synchronized ventricular depolarization. In particular, a non-null value of the time uniformity is expected since a value close to zero would mean an instantaneous depolarization through the whole ventricular myocardium that is unrealistic since a time lag of few milliseconds is known to occur between the activation of septal and apical myocytes.

3.3. Model reduction through sensitivity analysis

In this section, we aim to understand how sensitive the QoIs introduced above are on the model input parameters and

Table 3. Pearson and Spearman coefficients for both time and uniformity computed employing the 2500 independent samples of the 20 000 Saltelli's dataset.

input parameter	Pearson		Spearman	
	time	uniformity	time	uniformity
long axis	0.57	0.50	0.60	0.60
(SI) sphericity index	0.44	0.39	0.46	0.48
$M_{\text{int}}^{\parallel}$	−0.33	−0.29	−0.35	−0.35
$M_{\text{int}}^{\perp}$	−0.24	−0.20	−0.26	−0.24
$M_{\text{ext}}^{\parallel}$	−0.35	−0.31	0.34	−0.33
$M_{\text{ext}}^{\perp}$	−0.22	−0.20	−0.22	−0.22

detecting the most relevant ones. The dimension of the input parameter space will be then reduced through a variance-based analysis [50] using the first and total-order Sobol' indices as sensitivity indices [51]. The choice to use Sobol' indices is determined by the fact that the scatter plots in figure 5 do not exhibit a clear linear nor monotonic relation between the input and the QoIs, as also revealed by the low values of Pearson and Spearman coefficients [52] reported in table 3. Nevertheless, we anticipate that the Spearman coefficients will be a good predictor for the Sobol' indices ranking.

Figure 6. Graphical scheme of the adaptive strategy used to train and validate the optimal meta-model. Starting from the training dataset and using a least-squares error integration strategy, 64 metamodels are produced accordingly to a sequential/fixed selections strategy coupled with a linear/q-hyperbolic enumeration one and varying the polynomial degree from 1 to 16. To avoid overfitting, Q^2 indices are computed on an external independent test dataset and there are not any meta-models satisfying $Q^2 > 95\%$ the training dataset is increased, otherwise the R^2 coefficients of each metamodel passing the previous test is computed. Among the metamodel that also satisfy the underfitting test ($R^2 > 95\%$), the optimal metamodel is selected as the one minimizing $\|R^2 - Q^2\|$. In the case that any meta-models do not satisfy this last condition, the maximum total degree of the metamodel is increased and the procedure is restarted.

Given the computational cost of the electrophysiology model and the available computational resources, the number of possible model evaluations is more than what typically used to train a metamodel (dataset of the order of tens of samples), but anyway smaller than what would be needed for a direct strategy (dataset made by tens of thousands of samples). Therefore, we have decided to build a training and a testing dataset of a few hundred samples and train/test not one but a family of metamodels by varying the polynomial degree (ranging from 1 to 16), by using two different enumeration rules (linear and hyperbolic) and by adopting two different selection strategies (fixed and sequential). All the resulting $16 \times 2 \times 2 = 64$ metamodels are trained and validated against the corresponding dataset. The optimal one is selected among the family according to the method detailed below and sketched in figure 6. A different metamodel is obtained for each of the QoIs so that the trained metamodels are not only surrogate models of the initial complex electrophysiology system but, more importantly, are UQ tools for computing the corresponding global sensitivity indices needed to design a model reduction strategy. Furthermore, the computational cost of training another metamodel for a different QoI is negligible compared to the cost of building the dataset itself (see appendix A) and the UQ procedure allows for quick training of another metamodel in the case different QoIs are defined (in the order of a minute CPU time).

The first step is to produce a training dataset of 400 samples using a quasi-Monte Carlo (QMC) method (Sobol' sequence) so as to maximize the information contained and avoid samples clustering within the dataset [53]. The Sobol' sequence is used to generate the corresponding low discrepancy sequence of the samples [54]. The QMC has, indeed, a faster converge rate for a low number of parameters,

$\mathcal{O}\left(\frac{\log(N)^s}{N}\right)$ where s is the input dimension, compared to the standard MC, $\mathcal{O}\left(\frac{1}{\sqrt{N}}\right)$, [55]. Consequently, another 100 samples dataset independent of the training set is produced using a pure MC strategy, which will be used as a validation dataset. The cost of producing the whole dataset for the metamodel approach is approximately 17 CPU-days, where one CPU-day is defined as 1 day of computation done on a 1GHz reference processor.

Since truncated Gaussian distributions, as those used to model the input parameters, do not allow for analytical orthogonal polynomials with the respect to the norm weighted by the input PDFs, a family of orthogonal polynomials is produced numerically through the three terms recurrence [56]. It should be noted that, although the data from the literature allow estimation of the uncertainty PDFs of the input parameters, their statistical dependence/independence can not be determined. Hence, the input parameters are treated here as statistically independent and the multidimensional basis for the input space is then given by the product of the mono-dimensional basis of each input parameter. Given a computational model $G: \mathcal{D}_X \subset \mathbb{R}^d \rightarrow \mathbb{R}$, suppose that the uncertainty in the input parameters is modelled by a random vector X with prescribed joint probability density function $f_X(x)$ [57]. The resulting quantity of interest $Y = G(X)$ is obtained by propagating the uncertainty on X through G and, assuming that the input variables are statistically independent, the joint PDF is the product of the d marginal distributions $f_X(x) = \prod_{i=1}^d f_{X_i}(x_i)$ for each $\mathcal{D}_{X_i}$. For each single variable X_i and any two functions $\phi_1, \phi_2: \mathcal{D}_{X_i} \rightarrow \mathbb{R}$, we can define the inner product as $\langle \phi_1, \phi_2 \rangle := \int_{\mathcal{D}_{X_i}} \phi_1(x) \phi_2(x) f_{X_i}(x) dx$ and use it to define an orthonormal family of polynomials $\{P_k^{(i)}, k \in \mathbb{N}\}$. This set of univariate orthonormal polynomials can be used to define a family of multivariate ones.

In fact, given a multi-index $\alpha = (\alpha_1, \dots, \alpha_d)$, $\alpha_i \in \mathbb{N}$, the associated multivariate polynomial can be defined as $\Psi_\alpha(x) := \prod_{i=1}^d P_{\alpha_i}^{(i)}(x_i)$. The set of all multivariate polynomials in the input random vector X forms a basis of the Hilbert space [58], in which $Y = G(X)$ is given by the so-called polynomial chaos expansion

$$Y = \sum_{\alpha \in \mathbb{N}^d} y_\alpha \Psi_\alpha(X). \quad (3.3)$$

This infinite series has to be truncated in order to get a finite one approximating $Y = G(X)$ and different truncation strategies are possible depending (i) on how to enumerate the element of the multivariate basis and (ii) on how many terms of the basis have to be retained. We can define the standard, or linear enumeration strategy, based on the standard total degree of a multivariate polynomial Ψ_α , defined as $\|\Psi_\alpha\| := \sum_{i=1}^d \alpha_i$, that is the lexicographical order with a constraint of increasing total degree (e.g. for a two-dimensional multi-index $(0, 0) < (0, 1) < (1, 0) < (2, 0) < (1, 1) < \dots$).

However, the use of a standard enumeration is usually deprecated for biological applications because it yields oversized family of high-order interactions coefficients that are not observed in such cases [59]. A polynomial expansion coefficient is called a high-order interaction term when it is associated with a polynomial with high degree in more than one variable [60]. As physical phenomena are typically described by low-order interactions, an enumeration rule favouring low-order interactions rather than high-order ones is generally preferred [61]. With this motivation, as an alternative to the linear enumeration we have used a q -hyperbolic strategy that is based on the definition of a semi-norm characterizing each coefficient of the polynomial: given a real number $q \in (0, 1)$, the q -hyperbolic, quasi-norm of a multi-index α is defined as $\|\alpha\|_q = \left(\sum_{i=1}^d \alpha_i^q\right)^{1/q}$ and the space of the coefficients is explored (i.e. numerated) by increasing the value of the norm and selecting all the coefficients with a multi-index lower than the one selected. Smaller q yields a sharper hyperbolic selection and less high-order interactions (in terms of high degrees mixed coefficients of the PCE) for a fixed p .

Furthermore, two selection strategies have been selected: a *fixed strategy* where the total degree p is fixed and all the coefficients with norms (depending on the enumeration strategy) smaller or equal to p are retained, and the *sequential strategy* [61] that selects the most relevant coefficients in such a way to maximize the number of null coefficients.

For each metamodel corresponding to a different polynomial degree, truncation strategy and enumeration rule, the training dataset is used to evaluate its coefficients. The standard approach is to use high-order Gaussian quadrature rules providing the higher accuracy for a given number of samples [62], but this most accurate strategy has the drawback that the computational model has to be evaluated at the Gaussian integration points that depend on the accuracy of the integration formula. Moreover, in the case that the metamodel needs to be refined, the previous simulations can not be used and another dataset has to be produced from scratch. In order to circumvent these issues, the metamodel is fitted in a more flexible way using a least-squares error minimization of the coefficients [63] using the training dataset (400 samples).

In summary, we used a truncated PCE to approximate the model response. Firstly, a certain polynomial degree family (ranging from 1 to 16) is chosen, then the PCEs are truncated

by using two different enumeration strategies alternatively (linear and q -hyperbolic). Lastly, only a subset of the terms of each polynomial is retained according to the fixed or sequential truncation strategy for a total of $16 \times 2 \times 2 = 64$ possible metamodels that are trained on the given training dataset by minimizing the least-squares error. It should be noted that a linear enumeration coupled with a fixed truncation rule corresponds to a dimension of the basis with cardinality $\binom{d+p}{p}$, with p the polynomial degree and d the size of the input space [57].

Given this family of metamodels, the best one in terms of database fitting is then selected as follows. The risk of under/overfitting of the metamodels is assessed through a validation strategy commonly used in regression analysis, which is based on the coefficients R^2 and Q^2 [64]. The dataset is split into a training set that is used to calculate the index R^2 measuring the underfitting, and a validation set used to compute the index Q^2 measuring the overfitting. Given the training (testing) dataset of size n , described by the couples $(x_i, y_i)_{i=1}^n$ (where y_i is one of the QoIs corresponding to the set of input x_i) and the prediction of the metamodel f for the same input dataset $(x_i, \hat{y}_i := f(x_i))_{i=1}^n$, the coefficient of determination is defined as R^2 (Q^2) := $1 - (SS_r/SS_t)$ where $SS_r := \sum_{i=1}^n (y_i - \hat{y}_i)^2$ is the residual sum square normalized for the total sum of squares $SS_t := \sum_{i=1}^n (y_i - \bar{y})^2$ with $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$. The R^2 index is used in regression analysis to evaluate the goodness of the fit on the training dataset and a value close to 1 means that the metamodel correctly reproduces the variability within the training dataset whereas a lower R^2 index is a typical symptom of underfitting. On the other hand, a low Q^2 and high R^2 would correspond to an overfitting condition [64]. Importantly, the R^2 index does not measure the behaviour of the metamodel outside the training set and PCE is typically used in UQ analyses with few samples where all samples are used in the training set so as to avoid poor metamodeling [65]. As a consequence, the prediction error of the metamodel is typically evaluated through a leave-one-out cross-validation [66] that, however, does not provide such a consistent measure of the overfitting as done by the $R^2 - Q^2$ criterion used here [65]. The latter criterion, indeed, quantifies both the under- and overfitting with a large difference between $R^2 \approx 1$ and $Q^2 \ll 1$ implying that the metamodel has been over-adapted to the training dataset (overfitting) [67]. Hence, the optimal metamodel is chosen as the one minimizing the difference between these two coefficients, $\min \|R^2 - Q^2\|$ with the constraint that both parameters have to be larger than $\alpha = 95\%$, see figure 7a and b for the two QoIs of the first analysis. In the case none of the trained metamodels satisfies $R^2 > \alpha$, it implies that the polynomial family is not able to well approximate the training set (underfitting) and other metamodels with a higher number of coefficients should be trained. Conversely, if $Q^2 < \alpha$ for every metamodel it means that the validation set is poorly reproduced and the size of the training dataset has to be increased by adding samples. Importantly, the least-squares approximation of the PCE allows the dataset to be increased at any time, whereas a projection-based PCE needs to rebuild another dataset from scratch when a different size of the dataset is needed [61].

It turns out that the metamodels built using a sequential strategy are preferred to the ones corresponding to the fixed one, unless the total degree of the polynomial is low and the two selection strategies become equivalent selecting the same coefficients (and providing the same values for them).

Figure 7. R^2 (blue line) and Q^2 (orange line) coefficients as a function of the polynomial degree according to a hyperbolic enumeration for (a) the depolarization time with sequential truncation strategy and for (b) the depolarization uniformity with fixed truncation strategy. The vertical dashed lines indicate the total degree of the optimal metamodels that minimize the norm of $R^2 - Q^2$ (green lines, increased by a unit for visualization purposes).

Figure 8. Metamodel validation plot for the ventricular depolarization (a) time and (b) uniformity. Each point represents the QoI value computed using the electrophysiology model (abscissa) against the one evaluated by the metamodel (ordina) for a given set of input values. The green dots correspond to the output values of the training dataset (400 samples) while the red ones correspond to the results of the test dataset (100 samples). The solid blue line indicates an ideal perfect agreement between the metamodel and the full electrophysiology model.

In this last case, we indicate the optimal model as resulting from a fixed strategy to stress that all the coefficients have been selected. Hence, it turns out that the optimal metamodel depends on the specific QoI: a sequential truncation strategy with a total degree of 6 according to a hyperbolic enumeration ($q=0.4$) is found for the ventricular DT with $R^2=0.9769$ and $Q^2=0.97502$, whereas a fixed strategy with total degree 3 and hyperbolic enumeration ($q=0.4$) is the selected metamodel for the DU with $R^2=0.987$, $Q^2=0.979$. The corresponding fitting graph validating the optimal metamodel against the full electrophysiology model for the training (green dots) and test (red dots) datasets is reported in figure 8 for DT (a) and DU (b).

The metamodels can now be exploited to evaluate the sensitivity of the variance (Sobol' indices) of the QoIs to the input variables. Figure 9a indicates that the shape parameters L and SI influence over 60% of the Sobol' indices of the ventricular DT, whereas the perpendicular conductivities of the fibres have a cumulative influence of less than 15%. Note that as the ventricular shape parameters can be easily

measured *in vivo*, this is not the case for the electrical conductivities that are hardly measurable with common scanning techniques and relying on the results of the sensitivity analysis, the system can be reduced by fixing the two less relevant inputs (i.e. the intracellular and extracellular perpendicular conductivities) to their nominal values, thus decreasing the size of the input parameter space from 6 to 4. The same model reduction also applies to the ventricular DU, which exhibits a similar hierarchy of Sobol' indices in figure 9b and depends on the DT according to a linear law (Spearman coefficient of 0.998 and a Pearson coefficient of 0.992). We refer to appendix A for a convergence analysis of the Sobol' indices with respect to the size of the training dataset of the metamodel.

Before moving to the forward UQ analysis, a direct sampling approach, similar to the one proposed in [28], is used to validate the adaptive metamodel strategy. To this aim, a dataset (larger and independent of the 400+100 samples used for the metamodel approach) is built according to Saltelli's method [68,69] to compute the Sobol' indices (first and total order) using a pure MC sampling strategy

Figure 9. First- and total-order Sobol' indices for ventricular depolarization (a) time and (b) uniformity. Each index is computed using both a PCE strategy (sample size 400 + 100) and a direct one (sample size 20 000). The direct strategy allows to compute also the 95% Sobol' indices intervals here reported as error bars.

[70]. The size of the dataset $N_{ds} = 20\,000$ (instead of the 10 000 proposed in [28]) is chosen accordingly with the Saltelli size $N_{ds} = N \cdot (d + 2)$, where d is the number of parameters and N the number of independent entry, here $d = 6$ and $N = 2500$. This dataset requires a total of about 691 CPU-days to be produced. It should be remarked that the computational cost to produce the training + test dataset for the metamodel is about 2.5% of that needed for a direct strategy, while the cost to train the metamodel itself using the given methodology is of the order of 1 CPU-hour, therefore negligible compared to the dataset production. Figure 9 reports first- and total-order Sobol' indices obtained by the direct sampling technique superimposed to the ones evaluated by the metamodel. The comparison between the metamodel (trained on a 400+100 samples) and the direct validation strategy (calculated on 20 000 samples) shows as the metamodel predicts both the ranking of the first-order indices and their magnitude with a good accuracy (exception made for $M_e^{\parallel}$ for which the metamodel tends to underestimate the index). As expected, the direct approach appears to be less accurate on the total order, on which the confidence intervals are large (in particular for the DU) but, also in this case, the metamodel correctly identifies the relative order and approximates the indices well. On the other hand, the smaller confidence interval for the importance measures obtained from the asymptotic distribution of the Saltelli statistic, is a good indicator of the convergence of the algorithm. This comparison should be considered, we recall, only as a further check of the accuracy of the metamodels, as satisfying the $R^2 - Q^2$ criterion a sufficient validation. As a remark, the agreement between the two UQ techniques can probably be ascribed to the low-order nonlinearity of the problem that can be readily sampled by a small dataset.

Table 4. Statistical moments of the two QoIs according to the forward UQ analysis using the direct QMC sampling strategy (20 000 samples) on the reduced model.

	depolarization time	depolarization uniformity
mean	94.178 ms	22.90 ms
mean 95% confidence interval	[93.85, 94.51] ms	[22.82, 22.98] ms
standard deviation	23.86 ms	6.10 ms
skewness	1.14	1.01
kurtosis	5.28	5.04

3.4. Forward analysis

Once the sensitivity analysis has been performed, we can turn to the statistical characterization of the QoIs and evaluate their PDF using a direct UQ strategy. The sampling strategy is the QMC with Sobol's low discrepancy sequence (as detailed in the previous section) with a dataset made by 20 000 samples (≈ 691 CPU-days to be produced). Note that a direct strategy is the most robust method to evaluate the QoIs PDFs but, when this is not affordable owing to the large computational cost, the trained and validated metamodel is also used to produce the PDF best-fitting the data of the training set (see the analysis 2 in §4).

The computed stochastic moments are reported in table 4 showing for both QoIs a similar distribution shapes (also according to the skewness and the kurtosis) and a narrow 95% confidence interval on the mean. The corresponding PDFs of the QoIs are shown in figure 10a,b as histograms

Figure 10. PDF of the depolarization (a) time and (b) uniformity according to the forward analysis with QMC strategy on 20 000 cases. The grey histograms correspond to the computed model outputs (using Sylvester's bandwidth), whereas the red continuous distributions are lognormals (with parameters $\mu \log = 4.05$, $\sigma \log = 0.374$, $\gamma \text{ shift} = 32.23$ for (a) and $\mu \log = 3.25$, $\sigma \log = 0.221$, $\gamma \text{ shift} = -3.67$ for (b)).

(using Sylvester's rule for bandwidth [71]) with superimposed approximated lognormal distribution. The PDF of the ventricular DT in figure 10a can be read in light of the medical knowledge about the ventricular DT that is measured *in vivo* via ECG by monitoring the QRS complex (see figure 1c) and is considered normal as its value does not exceed 100 ms [1]. The UQ analysis predicts that, for input data corresponding to a healthy subject, the event of a normal DT happens with a probability of 65.8%, whereas a DT between 100 and 130 ms that is considered intermediate or slightly prolonged has a probability of 26.2%. Furthermore, a pathological DT longer than 130 ms has an 8% probability of occurring. Vice versa the probability of obtaining a DT shorter than 40 ms is lower than 10^{-6} because it is limited by the maximum speed of the electrical front propagation within the myocardium. The lognormal shape of the PDF for DT is probably related to the nonlinear relationship between the electrical conductivities and the conduction velocity and it was also obtained by Quaglino *et al.* [29] for the atrial myocardium by randomly perturbing both fibre orientation and local conductivities using a Bayesian multifidelity approach.

Owing to the quasi-linear relation between the two QoIs, the PDF of the DU (figure 10b) has a similar shape to that of the DT with a mean value of 6.10 ms. For this PDF, the probability of a variation of ± 10 ms with respect the mean value is equal to 8.1%, whereas it is very rare to observe a halving or doubling of the uniformity (still respect to its mean value) corresponding to a probability of 0.72%. This last result has important medical implications since a larger value of this QoI (as defined in (3.2)) corresponds to a less uniform ventricular depolarization and, as a consequence, a less effective blood ejection into the aorta during systole.

4. UQ analysis 2: sensitivity of the electrical activation of the left ventricle on the cellular model parameters

We now turn to investigate the UQ properties of the cellular model that determines the ionic current through the

membrane channels and which is coupled with the governing equations via the term I_{ion} in equations (2.1). With this analysis, we aim to investigate not only the sensitivity of the ventricular depolarization but also of the action potential profile on the input parameters of the ten Tusscher–Panfilov cellular model [30] and determine which of them are the most relevant. The ten Tusscher–Panfilov model, indeed, has been seen to well reproduce the kinematics of the ions and the corresponding action potential profile (see figure 1b), but its input parameters are usually taken equal to their nominal value by the user without exploring their sensitivity on the electrophysiology results, which will be studied in this section. Furthermore, we aim at understanding what parameters have to be varied in order to specialize the ten Tusscher–Panfilov equations to reproduce the action potential profile in pathological conditions or to model the myocardial depolarization in the heart chambers other than the LV (i.e. the atria and the right ventricle).

As recalled from introduction, this UQ analysis has been run separately with respect to the former (see §3) due to the lack of knowledge about the PDFs of the input parameters of the cellular model, which does not allow appropriate global sensitivity analysis to be performed [32]. Investigating simultaneously all the inputs would have polluted the resulting PDFs of the QoIs and, for this reason, two separated UQ analyses are considered rather than a single one on a larger input parameter space grouping the cellular model parameters with the geometrical and electrical input parameters. In this analysis, hence, the input parameters PDF are unknown and the UQ methodology is exploited to obtain more insight into the cellular model.

A possible approach to deal with input parameters with unknown uncertainty, would be to run a one-at-a-time sensitivity (OAT) analysis where each parameter p is varied around its mean μ within a range $\pm \alpha \mu$ in order to measure the sensitivity index $\frac{f[\mu(1+\alpha)]}{f[\mu(1-\alpha)]}$. The OAT analysis, however, has the main drawback that interactions among parameters would be neglected owing to the local variation of the parameters about their nominal values [31]. This approach has been extended here to run a semi-local analysis by varying each parameter uniformly across the entire range $[\mu(1-\alpha),$

Figure 11. QoIs for the analysis 2. (a) The action potential duration (APD) is the time in milliseconds occurring between the fast depolarization and the end of the repolarization phase. (b) Reference action potential (blue line) and a generic action potential (orange line) used to compute the shape similarity index.

Table 5. List of the input parameter of the cellular model considered in this analysis. Their nominal values are taken as in [30].

input parameter	definition	nominal value
G_{K1}	maximal IK1 conductance	5.405 nS pF^{-1}
G_{Kr}	maximal IKr conductance	0.153 nS pF^{-1}
G_{Ks}	maximal IKs conductance	0.392 nS pF^{-1}
G_{Na}	maximal INa conductance	$14.838 \text{ nS pF}^{-1}$
G_{bNa}	maximal IbNa conductance	$2.9 \times 10^{-4} \text{ nS pF}^{-1}$
G_{CaL}	maximal ICal conductance	$3.983 \times 10^{-5} \text{ nS pF}^{-1}$
G_{bCa}	maximal IbCa conductance	$5.92 \times 10^{-4} \text{ nS pF}^{-1}$
G_{to}	epicardial Ito conductance	0.294 nS pF^{-1}
G_{pCa}	maximal IpCa conductance	$0.1238 \text{ nS pF}^{-1}$
G_{pK}	maximal IpK conductance	$0.0146 \text{ nS pF}^{-1}$
Ca_o	extracellular Ca concentration	2 mM
Na_o	extracellular Na concentration	140 mM
K_o	extracellular K concentration	5.4 mM

$\mu(1 + \alpha)$ and using the dataset produced to calculate the Sobol' indices, which consider the parameters interactions. As this analysis is a generalization of the OAT sensitivity measure, it is natural to vary the inputs uniformly around their mean rather than using a tentative Gaussian distribution, which could yield an ill-posed global sensitivity analysis [52].

4.1. Input

The ten Tusscher–Panfilov cellular model represents the cellular-scale dynamics in the monodomain and bidomain equations and includes all the major ion channels as well as intracellular calcium dynamics. The resulting system of equations is made of 19 ODEs (not reported here for the sake of brevity) that depend on 53 input parameters. Some

of these input quantities, however, are constants, such as the gas and the Faraday constants, and some others are well controlled and measurable, such as the temperature and the surface to volume ratio of the cells. On the other hand, some other parameters are difficult to measure both *in vivo* and *ex vivo* and, despite significant variability among individuals is expected to happen, no data from the literature are available to estimate their PDFs. Therefore, among the 53 input parameters of the ten Tusscher–Panfilov cellular model, we focus here on the influence of the 10 conductance and three concentrations values of the ions, which are modelled as aleatory variables uniformly distributed in a range of $\pm 30\%$ around the nominal values reported in table 5. This variation range of the input parameters is larger than the one used in other parametric studies, where the effect of more parameters but varying in a smaller range was investigated [72].

4.2. QoIs

Along with the QoIs introduced in §3.2, i.e. the DT and DU, two more QoIs related to the action potential profile are considered here. The first one is the *action potential duration* (APD) defined as the time needed for a myocytes to return to the resting state and be reactive to another electrical stimulus. This quantity corresponds to the time interval occurring between the fast depolarization and the end of the slow repolarization of the myocytes (see figure 11a) and is measured in the numerical model as the time lag between the instant at which the cell reaches an active potential higher than -70 mV and the one at which the signal decreases lower than -80 mV (so as to avoid possible spurious depolarization signals triggered by small perturbations). The APD is measured using a fixed threshold instead of the APD_{90} (the time for 90% repolarization from maximum voltage), because the latter is defined using the action potential amplitude (APA), which is sensitive to the variation of the input parameters and could yield anomalous evaluation of the APD (see [72] and figure 16b–d).

As sketched in figure 11b, the second additional QoI is the *shape similarity index*, which measures the deviation of the action potential, $v(t)$, from the reference profile, $v_R(t)$,

Figure 12. Typical scatterplots of the four QoLs against some of the inputs parameters normalized by their nominal value reported in table 5 using the training dataset of the PCE (400 samples). The red boxes highlight possible input–QoLs relations that will be confirmed through a Sobol’ indices analysis.

Table 6. Selected metamodels for evaluating the different QoLs of analysis 2.

		total degree	enumeration strategy	truncation strategy	R^2	Q^2
QoL1	ventricular depolarization time	3	hyperbolic ($q = 0.4$)	fixed	0.995	0.995
QoL2	ventricular depolarization uniformity	10	hyperbolic ($q = 0.4$)	sequential	0.999	0.999
QoL3	action potential duration	12	hyperbolic ($q = 0.4$)	fixed	0.999	0.999
QoL4	action potential shape similarity index	7	hyperbolic ($q = 0.4$)	fixed	0.998	0.995

corresponding to the input parameters set to their nominal values (as reported in table 5). This QoI is defined as the maximum of the cross-correlation [73] between the action potential profile and the reference one normalized by the auto-correlation of $v_R^*(t)$ (i.e. the cross correlation of the reference profile with itself with time lag $\tau = 0$):

$$\text{shape similarity index } (v(t)) = \frac{\max_{\tau} \int_t v^*(t) v_R^*(t + \tau) dt}{\int_t v_R^*(t) v_R^*(t) dt}, \quad (4.1)$$

where $v^*(t) = v^*(t) + v_0$ and $v_R^*(t) = v_R^*(t) + v_0$ are the action potential profiles increased by the resting potential, $v_0 = -85$ mV. This definition of shape similarity index is commonly used in signal analysis to identify the starting point of a signal, even in the presence of noise [74]. Note that the formula (4.1) generally provides a shape similarity index less than one if $v(t)$ differs from the reference $v_R(t)$ with the two functions having a similar amplitude, whereas values greater than one occur when $v(t)$ has an amplitude larger than that of $v_R(t)$.

In order to compute the APD and the shape similarity index, we monitor the action potential at a sample node of the mesh that has to be far enough from the input signal in order to avoid an extra potential at the beginning of the electrical stimulus [11]. Since in a dedicated preliminary analysis the APD and the shape similarity index were seen to be

independent of the monitored myocardial location, without loss of generality both QoLs are here evaluated at the apex of the ventricle. These two additional QoLs investigated in this second analysis were not considered in §3 as a preliminary OAT investigation (not reported here for the sake of brevity) revealed that they are weakly sensitive to the electrical conductivities and to ventricular geometry.

4.3. Parametric study through sensitivity analysis

Owing to the large number of the input parameters, the QMC has an asymptotic slower convergence rate than the standard MC and about 10^4 samples are needed for the QMC to perform better than the MC with 13 input variables. For both the sampling methods, a direct strategy is not affordable in this case and the metamodel strategy introduced in §3 is applied by training the adaptive sparse PCE against a sample dataset of 1800 numerical simulations built using a Latin hypercube. An independent dataset of 200 samples is then created with a pure MC strategy to test the trained metamodels. The adaptive strategy described in §3 yields the metamodels reported in table 6. The cost of producing the entire dataset for training the metamodels is about 70 CPU-days, whereas the cost of a direct approach similar to the one used in the previous section

Figure 13. Spearman indices of analysis 2 computed using the training dataset (400 samples).

Figure 14. First- and total-order Sobol' indices computed using the adaptive PCE strategy on a dataset of 400 + 100 samples for (a) ventricular depolarization time and uniformity and (b) the APD and shape similarity index of the action potential.

would be about 3.5 CPU-years (with a Saltelli dimension of $2500 \times (13 + 2) = 37\,500$ samples).

The resulting scatter plots comparing the input variables and the corresponding QoIs are reported in figure 12 for some relevant cases: although, a clear input–QoIs correlation is not visible in all panels, some monotonic trends can be observed such as the maximal ICaL conductance, G_{CaL} , influencing the APD and the action potential shape similarity index or the maximal INa conductance, G_{Na} , and extracellular K concentration, K_o , affecting the ventricular DT and DU.

This monotonic behaviour is confirmed by the Spearman coefficients, which are good predictors for low order interactions between inputs and outputs and are reported in

figure 13 for the four QoIs against each input parameter. It can be noted that the DT and DU of the ventricle have a monotonic dependence on the maximal IbNa conductance, G_{bNa} , and extracellular K concentration, K_o , whereas the APD and the shape similarity index of the action potential manifest a monotonic dependence on the maximal IKs and ICaL conductances along with the extracellular K concentration, G_{Ks} , G_{CaL} and K_o , respectively. This preliminary statistical test thus confirms the low order input–output interactions that were suggested by the scatter plots in figure 12. Nevertheless, higher-order interaction among variables may be present as well and they need to be investigated before reducing the input parameters space. To this aim, the

Figure 15. Approximated polynomial response of the APD as a function of the four most relevant inputs (see text). The input parameter in abscissa have been normalized by the nominal value from the literature [30], which is indicated by the superscript ⁰.

importance measures and the Sobol' indices of the 13 input parameters of the cellular model are computed using the trained metamodels for each QoI. Figure 14*a* shows that only the maximal INa conductance, G_{Na} , the extracellular Na concentration, Na_o and the extracellular K concentration, K_o , influence the variance of the DT and DU, while these two QoIs are almost insensitive to the remaining 10 parameters, which have a total influence of less than 5% on the variance. Both these QoIs manifest a similar hierarchy of Sobol' indices that can be rationalized by computing their Pearson and Spearman coefficients that are both equal to 0.999. On the other hand, the importance measures and Sobol' indices of the other two QoIs (i.e. APD and shape similarity index) in figure 14*b* indicate that the most influential input parameters are the maximal IKr, IKs and ICaL conductances (G_{Kr} , G_{Ks} and G_{CaL}) along with the extracellular Ca and Na concentration (Ca_o , Na_o). Again, these two QoIs share the same behaviour of Sobol' indices with Pearson and Spearman coefficients of 0.940 and 0.947, respectively.

For a complete model reduction, the whole set of four QoIs must be taken into account noticing that six input parameters have a weak influence on any QoIs (namely G_{Kr} , G_{bNa} , G_{bCa} , G_{to} , G_{pCa} , G_{pK} and G_{Ks}) and the number of input variables can then be reduced from 13 to seven. A further reduction can then be done if only certain QoIs are chosen (e.g. the time and uniformity or shape and APD) as suggested by the Sobol' indices in figure 14.

In this analysis the difference between first order and total order indices is very little (in the order of 10^{-3}), which is in line with the choice of hyperbolic enumeration strategy that neglects high terms interaction, and allows a substantial reduction of the dimension of the samples needed to train the PCE. Furthermore, this low interaction among the input and the output variables permits the use of low order polynomials to better approximate (with a least-squares error of 10^{-4}) the QoI curves that are obtained by varying only one input parameter at the time while fixing the others to their nominal value, as shown in figure 15.

The sensitivity analysis carried out here thus provides the functional dependence between a local perturbation of the input parameter and the corresponding variation of the QoIs. This result can be applied to predict how the input parameters have to be set in such a way to reproduce action potential profiles occurring in pathological conditions or in the heart chambers other than the LV. For instance, Sobol'

indices suggest that the most effective parameters to increase or decrease the APD are G_{CaL} , G_{Ks} , Ca_o , Na_o and, as an example, these parameters have been varied in figure 16 showing how these values can be opportunely tuned to effectively modify and adapt the shape of the action potential. Vice versa a major modification of the other conductances and ions concentrations that are found to be irrelevant according to the sensitivity analysis do not yield a significant variation of the action potential (not reported in the figure).

4.4. Forward analysis

Although the lack of knowledge of the input PDFs makes the forward UQ not data based, the metamodels trained as in the previous subsection are here used to carry out a forward analysis by producing a large dataset (10^6 samples according to a pure MC scheme) that is used to evaluate the QoIs uncertainty. The computational cost of the forward analysis using a trained metamodel is negligible once the training is complete as the 20 000 metamodel evaluations needed for the UQ analysis are only obtained in about 1 CPU-minute, rather than 1.9 CPU-years as in the case of running the full computational model. The resulting PDFs are reported in figure 17 as histograms superimposed with the best-fitting distributions (obtained using the Kernel smoothing method). It results that varying the cellular model parameters within a range of $\pm 30\%$ over their nominal values does not produce pathological DTs since the standard deviation of its PDF is 9.2 ms and the probabilities for a DT greater than 120 ms or smaller than 60 ms are less than $10^{-2}\%$ and $10^{-10}\%$, respectively. Owing to the quasi-linear relation among the DT and DU observed in §3.3, the distribution of the ventricular DU is concentrated about its mean value with a standard deviation equal to 2.31 ms and a probability of below 2.5% to experience a DU shifted by more than 5 ms from its mean. Interestingly, the asymmetric distribution of the DT and DU reported in figure 17*a* and *b* is due to the nonlinear influence of the cellular model on the conduction velocity that has a superlinear (sublinear) dependence on the electrical conductivities for small (large) values of the electrical conductivities themselves.

On the other hand, the APD exhibits a more symmetric PDF (skewness 0.05) with a mean of 295.8 ms and a standard deviation equal to 23.4 ms (see figure 17*c*). The probability of a APD 30 ms larger than the mean value, corresponding to a

Figure 16. Variation of the action potential corresponding to doubling and halving the most sensitive input parameter according to the sensitivity analysis: (a) Ca_O , (b) G_{CaL} , (c) G_{Ks} and (d) Na_O .

Figure 17. PDF of the depolarization (a) time and (b) uniformity, (c) APD and (d) shape similarity index of the action potential according to the forward analysis. The grey histograms correspond to the computed metamodel outputs.

longer inability of the muscular fibre to contract, is equal to 20.8%. Lastly, the PDF of the action potential similarity index has an average of 0.98 and a standard deviation of 4×10^{-3} where, we recall, the unitary value corresponds to the reference action potential shape. The asymmetric distribution with a skewness of -0.58 and the absence of a right tail means that the main effect on the action potential shape is a modification with respect to the reference profile (yielding indeed a cross-correlation less than one according to the equation (4.1)), rather than a variation of its amplitude corresponding to a shape similarity index greater than one.

5. Conclusion

In this work, the UQ analysis for an electrophysiology model of the human LV has been performed. The first analysis is focused on the effect of the ventricular geometry and electrical conductivities (input parameters) on the effectiveness of ventricular contraction during the systolic phase that is known to be correlated with (i) the time needed to depolarize the whole myocardium and (ii) the spatial uniformity of the depolarization front (QoIs). The uncertain PDFs of the shape parameters are taken from available data in the literature acquired with echocardiography, whereas the electrical conductivities are known to be distributed as Gaussians for mammals and the few available measurements for humans are used to calibrate Gaussian PDFs. Thereafter, a sensitivity analysis has been carried out using an adaptive strategy based on a training dataset (400 samples) produced using a QMC strategy (with Sobol' low discrepancy sequence) and an independent test dataset (100 samples) obtained using a pure MC strategy. A family of metamodels (PCE based on orthogonal polynomials recovered using a three terms recursion) has been built by using (i) two different selection strategies of their coefficients (fixed or sequential), (ii) two enumeration rules (linear or hyperbolic) and varying (iii) the total maximum degree of the expansion from 1 to 16. Each of the 64 resulting metamodels is trained against the training dataset and the optimal one is selected as the one minimizing the distance between the R^2 index (evaluated on the training dataset) and the Q^2 index (evaluated on the test dataset), with the constraint $R^2, Q^2 > 95\%$. The optimal metamodel is then used to compute the first and total Sobol' indices to determine the QoI sensitivities on the input parameters. In order to validate the sensitivity analysis, a direct *brute force* approach has also been applied on an independent QMC dataset of 20 000 samples to produce Sobol' indices using Saltelli's algorithm that well agree with the one obtained using the metamodeling approach. PCE metamodels, indeed, are seen to fit several electrophysiology models well and to provide sensitivity results in line with direct strategy approaches but using smaller datasets than the ones needed for a direct UQ analysis (see figure 9 for an example) [28]. This high prediction ability, combined with the stability of electrophysiology models at small perturbations, places PCE strategies among the gold standards techniques for electrophysiological UQ [72].

The analysis reveals that ventricular depolarization is very sensitive to the geometrical parameters (with an influence greater than 60%), while the parallel and perpendicular conductivities have a total influence of about 25% and 15% respectively. This result thus suggests reducing the

model by neglecting the perpendicular conductivities and decreasing the size of the input parameter space from six to four. Interestingly, a strong linear correlation between the DT and DU is observed (Pearson index 0.992 and Spearman index 0.998) implying that the same model reduction applies to both QoIs. The reduced model is then used to run a forward analysis (QMC dataset with Sobol' low discrepancy sequence) and obtain the uncertainty PDFs that are seen to be in-line with clinical observations exhibiting a non-pathologic DT with a probability of 80%, whereas a slow DT longer than 130 ms has a probability of about 8%. On the other hand, the probability of unphysically low DT (i.e. less than 40 ms) is smaller than 10^{-6} .

We then turned our attention to investigate the role of the input parameters of the ten Tusscher–Panfilov cellular model on the DT and DU of the LV along with two more QoIs related to the active potential profile: the APD and the shape similarity index (the latter defined with respect to the reference case corresponding to the input parameters set to their nominal values). This second analysis has been run separately from the first since no data from the literature are available to estimate the input uncertainties, which have been modelled as independent uniform distributions with values ranging from 0.7 to 1.3 times the nominal value. In this parametric study, a direct UQ strategy is unaffordable owing to the large number of parameters involved (10 conductances plus the three ion concentrations), and the adaptive metamodel strategy used in the first analysis has been adapted to this case to run a sensitivity analysis. The resulting Sobol' indices show that the DT and DU only depend on the maximal INa conductance, G_{Na} , extracellular Na concentration, Na_o , and extracellular K concentration, K_o , whereas the other 10 input parameters have a total influence of less than 5% on the variance. The action potential QoIs (i.e. APD and shape similarity index) are sensitive to the maximal conductances G_{CaL} , G_{Ks} and G_{Kr} , along with the extracellular Ca and Na concentrations Na_o and G_{Kr} , with a total influence exceeding 95%. This result naturally suggests a model reduction from 12 to five input variables, namely G_{CaL} , G_{Ks} , Ca_o , Na_o and G_{Kr} .

A similar strategy for the model reduction was obtained by Hurtado *et al.* [27] where a PCE approach based on the R^2 validation index, similar to the one adopted in our analysis, is used. Although these authors considered different cellular models compared to ours (the Nash–Panfilov and the Rice models) and a smaller dimension of the input parameters space (five independent inputs, rather than 12 investigated here), the same order of importance for the four common input parameters on the APD is observed (namely G_{CaL} , G_{Ks} , G_{Kr} and G_{Na} with a negligible effect of the latter one on the APD). Furthermore, the sign of the APD sensitivity on each input parameter reported in Hurtado *et al.* corresponds to those computed here using the Spearman's indices (see figure 13). The sensitivity analysis and the corresponding model reduction relative to the APD is also confirmed by Pathmanathan and Cordeiro [72,75], where a parametric study with normal input distributions was carried out using a direct UQ approach on several QoIs. Also in their work G_{Kr} , G_{Ks} and G_{CaL} are the relevant input parameters on the APD that, instead, is less sensitive on G_{K1} , G_{Na} and G_{to} .

On the other hand, if the whole set of four QoIs is considered, six variables among 13 are seen to have a weak influence on any QoI and they can be neglected with the aim of reducing

the model's complexity. Importantly, the difference between first- and total-order indices is small (in the order of 10^{-3}) thus meaning that the cellular model inputs could be varied separately to modify the QoIs. This result has then been exploited to successfully control the APD of the cellular model by perturbing the sensitive input parameters detected by the sensitivity analysis. This analysis thus provides a deeper insight into the effects of the input parameters on the action potential shape and the sensitive input parameters can be used in future studies to reproduce the transmembrane action potential heterogeneity in the heart chambers and/or reproduce the action potential profiles observed in pathological conditions.

A main limitation of this work, however, is that the electrophysiology model solves for the LV as uncoupled from the rest of the heart chambers by setting an initial electrical stimulus at the location of the bundle of His, thus neglecting the electrical depolarization of the atria that is initiated in the SA-node. Furthermore, the electrical system of fast conduction of the heart including the highly conductive bundles and the Purkinje network are also neglected and a more evolute and accurate electrophysiology model is needed to account for these limitations.

Nevertheless, this work represents a starting point to detect the key input parameters influencing the electrical activation of the ventricular myocardium, to quantify the uncertainty on the depolarization pattern and to reduce the input parameter space. This last aspect is of paramount importance when the computational cost of the model is further increased as in the case of multi-physics models for the heart where such an electrophysiology model is coupled with the structural dynamics of the myocardium (electro-mechanical interaction) and the mechanical stresses influence the conductivity properties of cardiac tissue through a generalized reaction–diffusion–mechanics model [76,77]. Such refined computational models introduce additional input parameters and QoIs, thus calling for further UQ analyses. Future UQ analyses could thus be carried out by considering other relevant quantities to describe critical features in the cardiac dynamics such as the electrical conduction velocity or, in the case of electro-fluid–structure interactions (FSEI [18]), fluid-dynamics QoIs such as recirculation zones, flow patterns or the cardiac output. The size of the input parameter space could also be increased by including the uncertainty of the fibre orientation, of the specific membrane capacitance C_m and of the surface-to-volume ratio of cells χ as well as by considering different positions and patterns of the stimulus I_s . Indeed, at higher stimulation frequencies, other ionic parameters can become relevant [78,79] and *ad hoc* QoIs characterizing the action potential patterns can be introduced [80].

As a final note, we wish to remark that the accuracy of the UQ analyses proposed here as further development of this work depend on the uncertainty PDFs of the input parameters that need to be retrieved from real data through dedicated *in vitro* and *in vivo* laboratory experiments.

Data accessibility. This article has no additional data.

Authors' contributions. G.D.C., R.V., F.V. designed the research. G.D.C. carried out the numerical simulations and the UQ analysis. G.D.C. and F.V. wrote the paper with inputs from R.V.

Competing interests. We declare we have no competing interests.

Funding. This work has been partly supported by grant no. 2017A889FP 'Fluid dynamics of hearts at risk of failure: towards methods for the prediction of disease progressions' funded by the Italian Ministry of Education and University.

Appendix A. Convergence analyses

A.1. Convergence check of the electrophysiology model

The transmembrane potential averaged over the ventricular domain as a function of time is shown in figure 18*a*. Each solid curve corresponds to a different simulation of the bidomain equations with the ten Tusscher–Panfilov model on a different grid (with cells number varying from 2000 to 68 000), where the electrical conductivities are scaled by a fixed parameter equal to 3.0 to account for the lack of fast conductivity fibres (see §2). The resulting transient dynamics of the average transmembrane potential evolves in a similar fashion to a typical action potential profile with a steep depolarization front and a smoother repolarization one occurring after about 300 ms (figure 1*b*). Qualitative differences of this quantity are visible when the number of cells is below 3000 cells, whereas the results become insensitive on the grid resolution of over 50 000 cells. Such a resolution, however, would require a CPU time exceeding 20 000 days to build a dataset as large as the one used for the direct UQ analysis in §3 (20 000 samples), thus calling for a massively parallel CPU infrastructure. According to dedicated numerical tests, the computational cost to solve a single time step of the bidomain model is found to increase linearly with the number of mesh cells and, at the same time, the largest time step ensuring numerical stability is seen to decrease about with the same scaling. As a result, the total computational cost has a quadratic dependence on the number of mesh cells as shown in figure 18*b*. In order to reduce the computational cost to match the computational resources available, we have conveniently used the mesh with 4311 cells and scaled the electrical conductivities by a fixed parameter equal to 4.99 accounting both for the lack of the fast conductivity fibres and for the correction of the conduction velocities to mesh size effects. This method is commonly used in modomain/bidomain modelling [81–83] as it captures a more accurate approximation of the conductive velocity (see the dashed line in figure 18*a*) and, consequently, can build a 20 000 samples dataset with a CPU cost of 691 days.

A.2. Convergence check of the metamodel UQ method

The accuracy of both the metamodel and the direct UQ approaches has been already shown by the cross-validation test reported in §3.3, where the first and total-order Sobol' indices obtained with the two techniques agree well with each other, see figure 9. A convergence study of the metamodel approach is carried out here by varying the size of the training dataset. To this aim, smaller training datasets are obtained by extracting a subset of the training dataset used in §3.3 containing from 50 to 400 samples with a step of 50 (in the last case, the subset corresponds to the whole training dataset) and for each case, a different metamodel is trained using the same adaptive technique introduced above. Figure 19 shows the first-order Sobol' indices as a function of the number of samples of the dataset used to train the metamodel: as the dataset is larger than 200 samples (i.e. half of the dataset used for training the metamodel in §3.3), the measures of importance are almost stable thus guaranteeing the convergence of the method. It should be noted that this good behaviour of the metamodel is also attributable to the sampling procedure (QMC), which maps the input space effectively identifying the average properties with few samples.

Figure 18. (a) Time behaviour of the average transmembrane potential in the left ventricle with electrical conductivities scaled by a fixed parameter equal to 3. The dashed line indicates the configuration used for the UQ analysis, corresponding to the 4311 cells grid with electrical conductivities scaled by a factor 4.99. (b) Computational cost in CPU days of producing a dataset of 20 000 samples (as large as the one used for the direct UQ analysis in §3) as a function of the mesh cells number. The vertical dashed line indicates the computational cost corresponding to the configuration used to the UQ analyses in §§3 and 4.

Figure 19. Convergence analysis for the metamodel training procedure. First-order Sobol' indices of analysis 1 (see §3) as a function of the size of the training dataset for (a) DT and (b) DU.

Figure 20. Convergence analysis for the direct UQ procedure. First-order Sobol' indices of analysis 1 (see §3) using the smaller (4000 samples) and larger (20 000 samples) datasets.

A.3. Convergence check of the direct UQ method

The convergence of the direct QMC method is tested by applying the same QMC UQ procedure used for analysis 1 in §3.3 to

an independent and smaller Saltelli dataset of size 4000 samples. Figure 20 shows that the first-order Sobol' indices computed using the two datasets reasonably agree each other

and, importantly, manifest the same order of importance (except for the SI and $M_{\text{ext}}^{\parallel}$ on the DU) despite that this second dataset is five times smaller than the one used in §3.3.

Furthermore, the importance measures of DT are within the confidence intervals calculated on the larger Saltelli dataset of 20 000 samples.

References

- Hall JE. 2015 *Guyton and Hall textbook of medical physiology e-Book*. Philadelphia, PA: Elsevier Health Sciences.
- Anderson KR, Ho SY, Anderson RH. 1979 Location and vascular supply of sinus node in human heart. *Heart* **41**, 28–32. (doi:10.1136/hrt.41.1.28)
- Jack JJB, Noble D, Tsien RW. 1975 *Electric current flow in excitable cells*. Oxford, UK: Clarendon Press.
- Leon L, Roberge F. 1991 Directional characteristics of action potential propagation in cardiac muscle. A model study. *Circ. Res.* **69**, 378–395. (doi:10.1161/01.RES.69.2.378)
- Vigmond EJ, Hughes M, Plank G, Leon LJ. 2003 Computational tools for modeling electrical activity in cardiac tissue. *J. Electrocardiol.* **36**, 69–74. (doi:10.1016/j.jelectrocard.2003.09.017)
- Bueno-Orovio A, Kay D, Grau V, Rodriguez B, Burrage K. 2014 Fractional diffusion models of cardiac electrical propagation: role of structural heterogeneity in dispersion of repolarization. *J. R. Soc. Interface* **11**, 20140352. (doi:10.1098/rsif.2014.0352)
- Hurtado DE, Castro S, Gizzi A. 2016 Computational modeling of non-linear diffusion in cardiac electrophysiology: a novel porous-medium approach. *Comput. Methods Appl. Mech. Eng.* **300**, 70–83. (doi:10.1016/j.cma.2015.11.014)
- Pullan AJ, Tomlinson KA, Hunter PJ. 2002 A finite element method for an eikonal equation model of myocardial excitation wavefront propagation. *SIAM J. Appl. Math.* **63**, 324–350. (doi:10.1137/S0036139901389513)
- Sepulveda NG, Roth BJ, Wikswo Jr JP. 1989 Current injection into a two-dimensional anisotropic bidomain. *Biophys. J.* **55**, 987. (doi:10.1016/S0006-3495(89)82897-8)
- Tung L. 1978 *A bi-domain model for describing ischemic myocardial dc potentials*. Cambridge, MA: MIT.
- Sundnes J, Lines GT, Cai X, Nielsen BF, Mardal KA, Tveito A. 2007 *Computing the electrical activity in the heart*, vol. 1. Berlin/Heidelberg, Germany: Springer Science & Business Media.
- Potse M, Dubé B, Richer J, Vinet A, Gulrajani RM. 2006 A comparison of monodomain and bidomain reaction-diffusion models for action potential propagation in the human heart. *IEEE Trans. Biomed. Eng.* **53**, 2425–2435. (doi:10.1109/TBME.2006.880875)
- Vigmond E, Dos Santos RW, Prassl A, Deo M, Plank G. 2008 Solvers for the cardiac bidomain equations. *Prog. Biophys. Mol. Biol.* **96**, 3–18. (doi:10.1016/j.pbiomolbio.2007.07.012)
- Roth BJ. 2001 Meandering of spiral waves in anisotropic cardiac tissue. *Physica D* **150**, 127–136. (doi:10.1016/S0167-2789(01)00145-2)
- Trayanova N. 2006 Defibrillation of the heart: insights into mechanisms from modelling studies. *Exp. Physiol.* **91**, 323–337. (doi:10.1113/expphysiol.2005.030973)
- Wikswo Jr JP, Lin SF, Abbas RA. 1995 Virtual electrodes in cardiac tissue: a common mechanism for anodal and cathodal stimulation. *Biophys. J.* **69**, 2195–2210. (doi:10.1016/S0006-3495(95)80115-3)
- Muzikant A, Henriquez C. 1998 Validation of three-dimensional conduction models using experimental mapping: are we getting closer? *Prog. Biophys. Mol. Biol.* **69**, 205–223. (doi:10.1016/S0079-6107(98)00008-X)
- Viola F, Meschini V, Verzicco R. 2020 Fluid–Structure–Electrophysiology interaction (FSEI) in the left-heart: a multi-way coupled computational model. *Eur. J. Mech.-B/Fluids* **79**, 212–232. (doi:10.1016/j.euromechflu.2019.09.006)
- Britton OJ, Bueno-Orovio A, Van Ammel K, Lu HR, Towart R, Gallacher DJ, Rodriguez B. 2013 Experimentally calibrated population of models predicts and explains intersubject variability in cardiac cellular electrophysiology. *Proc. Natl Acad. Sci. USA* **110**, E2098–E2105. (doi:10.1073/pnas.1304382110)
- Sermesant M *et al.* 2012 Patient-specific electromechanical models of the heart for the prediction of pacing acute effects in CRT: a preliminary clinical validation. *Med. Image Anal.* **16**, 201–215. (doi:10.1016/j.media.2011.07.003)
- Johnston PR. 2011 A sensitivity study of conductivity values in the passive bidomain equation. *Math. Biosci.* **232**, 142–150. (doi:10.1016/j.mbs.2011.05.004)
- Barone A, Gizzi A, Fenton F, Filippi S, Veneziani A. 2020 Experimental validation of a variational data assimilation procedure for estimating space-dependent cardiac conductivities. *Comput. Methods Appl. Mech. Eng.* **358**, 112615. (doi:10.1016/j.cma.2019.112615)
- Cusimano N, Gizzi A, Fenton F, Filippi S, Gerardo-Giorda L. 2020 Key aspects for effective mathematical modelling of fractional-diffusion in cardiac electrophysiology: a quantitative study. *Commun. Nonlinear Sci. Numer. Simul.* **84**, 105152. (doi:10.1016/j.cnsns.2019.105152)
- Yang H, Veneziani A. 2015 Estimation of cardiac conductivities in ventricular tissue by a variational approach. *Inverse Prob.* **31**, 115001. (doi:10.1088/0266-5611/31/11/115001)
- Konukoglu E *et al.* 2011 Efficient probabilistic model personalization integrating uncertainty on data and parameters: application to eikonal-diffusion models in cardiac electrophysiology. *Prog. Biophys. Mol. Biol.* **107**, 134–146. (doi:10.1016/j.pbiomolbio.2011.07.002)
- Chan K, Saltelli A, Tarantola S. 1997 Sensitivity analysis of model output: variance-based methods make the difference. In *Proc. of the 29th Conf. on Winter simulation*, pp. 261–268.
- Hurtado DE, Castro S, Madrid P. 2017 Uncertainty quantification of 2 models of cardiac electromechanics. *Int. J. Numer. Methods Biomed. Eng.* **33**, e2894. (doi:10.1002/cnm.2894)
- Hu Z, Du D, Du Y. 2018 Generalized polynomial chaos-based uncertainty quantification and propagation in multi-scale modeling of cardiac electrophysiology. *Comput. Biol. Med.* **102**, 57–74. (doi:10.1016/j.combiomed.2018.09.006)
- Quaglino A, Pezzuto S, Koutsourelakis PS, Auricchio A, Krause R. 2018 Fast uncertainty quantification of activation sequences in patient-specific cardiac electrophysiology meeting clinical time constraints. *Int. J. Numer. Methods Biomed. Eng.* **34**, e2985. (doi:10.1002/cnm.2985)
- Ten Tusscher K, Panfilov A. 2006 Cell model for efficient simulation of wave propagation in human ventricular tissue under normal and pathological conditions. *Phys. Med. Biol.* **51**, 6141. (doi:10.1088/0031-9155/51/23/014)
- Hamby D. 1995 A comparison of sensitivity analysis techniques. *Health Phys.* **68**, 195–204. (doi:10.1097/00004032-199502000-00005)
- Iooss B, Lemaitre P. 2015 A review on global sensitivity analysis methods. In *Uncertainty management in simulation-optimization of complex systems*, pp. 101–122. New York, NY: Springer.
- Niederer SA *et al.* 2011 Verification of cardiac tissue electrophysiology simulators using an N-version benchmark. *Phil. Trans. R. Soc. A* **369**, 4331–4351. (doi:10.1098/rsta.2011.0139)
- E Rognes M, E Farrell P, W Funke S, E Hake J, MC Maleckar M. 2017 cbcbeat: an adjoint-enabled framework for computational cardiac electrophysiology. *J. Open Sour. Softw.* **2**, 224. (doi:10.21105/joss.00224)
- Alnæs M *et al.* 2015 The FEniCS Project Version 1.5. *Arch. Numer. Softw.* **3**, 163–202.
- Logg A, Mardal KA, Wells GN eds. 2012 *Automated solution of differential equations by the finite element method*. Berlin/Heidelberg, Germany: Springer.
- Chapelle D, Collin A, Gerbeau JF. 2013 A surface-based electrophysiology model relying on asymptotic analysis and motivated by cardiac atria modeling. *Math. Models Methods Appl. Sci.* **23**, 2749–2776. (doi:10.1142/S0218202513500450)
- Collin A, Gerbeau JF, Hocini M, Haïssaguerre M, Chapelle D. 2013 Surface-based electrophysiology modeling and assessment of physiological

- simulations in atria. In *Int. Conf. on Functional Imaging and Modeling of the Heart*, pp. 352–359. New York, NY: Springer.
39. Deshmukh P, Casavant DA, Romanynshyn M, Anderson K. 2000 Permanent, direct His-bundle pacing: a novel approach to cardiac pacing in patients with normal His–Purkinje activation. *Circulation* **101**, 869–877. (doi:10.1161/01.CIR.101.8.869)
40. Di Donato M *et al.* 2006 Left ventricular geometry in normal and post-anterior myocardial infarction patients: sphericity index and new concity index comparisons. *Eur. J. Cardiothorac. Surg.* **29**(suppl. 1), S225–S230. (doi:10.1016/j.ejcts.2006.03.002)
41. Tarantola A. 2005 *Inverse problem theory and methods for model parameter estimation*, vol. 89. Philadelphia, PA: SIAM.
42. Greiner J, Sankarankutty AC, Seemann G, Seidel T, Sachse FB. 2018 Confocal microscopy-based estimation of parameters for computational modeling of electrical conduction in the normal and infarcted heart. *Front. Physiol.* **9**, 239. (doi:10.3389/fphys.2018.00239)
43. Clerc L. 1976 Directional differences of impulse spread in trabecular muscle from mammalian heart. *J. Physiol.* **255**, 335–346. (doi:10.1113/jphysiol.1976.sp011283)
44. Roberts DE, Scher AM. 1982 Effect of tissue anisotropy on extracellular potential fields in canine myocardium *in situ*. *Circ. Res.* **50**, 342–351. (doi:10.1161/01.RES.50.3.342)
45. Roberts DE, Hersh LT, Scher AM. 1979 Influence of cardiac fiber orientation on wavefront voltage, conduction velocity, and tissue resistivity in the dog. *Circ. Res.* **44**, 701–712. (doi:10.1161/01.RES.44.5.701)
46. Roth BJ. 1988 The electrical potential produced by a strand of cardiac muscle: a bidomain analysis. *Ann. Biomed. Eng.* **16**, 609–637. (doi:10.1007/BF02368018)
47. Roth BJ. 1997 Nonsustained reentry following successive stimulation of cardiac tissue through a unipolar electrode. *J. Cardiovasc. Electrophysiol.* **8**, 768–778. (doi:10.1111/j.1540-8167.1997.tb00835.x)
48. Massey Jr FJ. 1951 The Kolmogorov–Smirnov test for goodness of fit. *J. Am. Stat. Assoc.* **46**, 68–78. (doi:10.1080/01621459.1951.10500769)
49. Dowson D, Wragg A. 1973 Maximum-entropy distributions having prescribed first and second moments (corresp.). *IEEE Trans. Inf. Theory* **19**, 689–693. (doi:10.1109/TIT.1973.1055060)
50. Chan K, Tarantola S, Saltelli A, Sobol IM. 2000 Variance-based methods. *Sensitivity Anal.* **45**, 167–197.
51. Homma T, Saltelli A. 1996 Importance measures in global sensitivity analysis of nonlinear models. *Reliab. Eng. Syst. Saf.* **52**, 1–17. (doi:10.1016/0951-8320(96)00002-6)
52. Saltelli A *et al.* 2000 *Sensitivity analysis*. Probability and statistics series. New York, NY: John and Wiley & Sons.
53. Owen AB. 2000 Monte Carlo, quasi-Monte Carlo, and randomized quasi-Monte Carlo. In *Monte-Carlo and quasi-Monte Carlo methods*, pp. 86–97. New York, NY: Springer.
54. Sobol' IM. 1967 On the distribution of points in a cube and the approximate evaluation of integrals. *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki* **7**, 784–802.
55. Caflisch RE. 1998 Monte Carlo and quasi-Monte Carlo methods. *Acta Numer.* **7**, 1–49. (doi:10.1017/S0962492900002804)
56. Gautschi W. 2004 *Orthogonal polynomials*. Oxford, UK: Oxford university press.
57. Gratiot LL, Marelli S, Sudret B. 2016 Metamodel-based sensitivity analysis: polynomial chaos expansions and Gaussian processes. In *Handbook of Uncertainty Quantification*, pp. 1–37.
58. Soize C, Ghanem R. 2004 Physical systems with random uncertainties: chaos representations with arbitrary probability measure. *SIAM J. Sci. Comput.* **26**, 395–410. (doi:10.1137/S1064827503424505)
59. C Montgomery D. 1997 *Montgomery design and analysis of experiments*. Hoboken, NJ: John Wiley.
60. Blatman G, Sudret B. 2010 Efficient computation of global sensitivity indices using sparse polynomial chaos expansions. *Reliab. Eng. Syst. Saf.* **95**, 1216–1229. (doi:10.1016/j.res.2010.06.015)
61. Sudret B. 2008 Global sensitivity analysis using polynomial chaos expansions. *Reliab. Eng. Syst. Saf.* **93**, 964–979. (doi:10.1016/j.res.2007.04.002)
62. Eldred M, Burkardt J. 2009 Comparison of non-intrusive polynomial chaos and stochastic collocation methods for uncertainty quantification. In *47th AIAA aerospace sciences meeting including the new horizons forum and aerospace exposition*, p. 976.
63. Levenberg K. 1944 A method for the solution of certain non-linear problems in least squares. *Q. Appl. Math.* **2**, 164–168. (doi:10.1090/qam/10666)
64. Consonni V, Ballabio D, Todeschini R. 2010 Evaluation of model predictive ability by external validation techniques. *J. Chemom.* **24**, 194–201. (doi:10.1002/cem.1290)
65. Anderssen E, Dyrstad K, Westad F, Martens H. 2006 Reducing over-optimism in variable selection by cross-model validation. *Chemom. Intell. Lab. Syst.* **84**, 69–74. (doi:10.1016/j.chemolab.2006.04.021)
66. Wu CFJ *et al.* 1986 Jackknife, bootstrap and other resampling methods in regression analysis. *Ann. Stat.* **14**, 1261–1295. (doi:10.1214/aos/1176350142)
67. Mohri M, Rostamizadeh A, Talwalkar A. 2018 *Foundations of machine learning*. Cambridge, MA: MIT Press.
68. Saltelli A. 2002 Making best use of model evaluations to compute sensitivity indices. *Comput. Phys. Commun.* **145**, 280–297. (doi:10.1016/S0010-4655(02)00280-1)
69. Smith RC. 2013 *Uncertainty quantification: theory, implementation, and applications*, vol. 12. Philadelphia, PA: Siam.
70. Metropolis N, Ulam S. 1949 The Monte Carlo method. *J. Am. Stat. Assoc.* **44**, 335–341. (doi:10.1080/01621459.1949.10483310)
71. Scott DW. 2015 *Multivariate density estimation: theory, practice, and visualization*. Hoboken, NJ: John Wiley & Sons.
72. Pathmanathan P, Cordeiro JM, Gray RA. 2019 Comprehensive uncertainty quantification and sensitivity analysis for cardiac action potential models. *Front. Physiol.* **10**, 721. (doi:10.3389/fphys.2019.00721)
73. Martin J, Crowley JL. 1995 Comparison of correlation techniques. In *Intelligent Autonomous Systems*, pp. 86–93.
74. Coakley KJ, Hale P. 2001 Alignment of noisy signals. *IEEE Trans. Instrum. Meas.* **50**, 141–149. (doi:10.1109/19.903892)
75. Pathmanathan P, Shotwell MS, Gavaghan DJ, Cordeiro JM, Gray RA. 2015 Uncertainty quantification of fast sodium current steady-state inactivation for multi-scale models of cardiac electrophysiology. *Prog. Biophys. Mol. Biol.* **117**, 4–18. (doi:10.1016/j.pbiomolbio.2015.01.008)
76. Cherubini C, Filippi S, Gizzi A, Ruiz-Baier R. 2017 A note on stress-driven anisotropic diffusion and its role in active deformable media. *J. Theor. Biol.* **430**, 221–228. (doi:10.1016/j.jtbi.2017.07.013)
77. Loppini A, Gizzi A, Ruiz-Baier R, Cherubini C, Fenton FH, Filippi S. 2018 Competing mechanisms of stress-assisted diffusivity and stretch-activated currents in cardiac electromechanics. *Front. Physiol.* **9**, 1714. (doi:10.3389/fphys.2018.01714)
78. Gizzi A, Cherry E, Gilmour Jr RF, Luther S, Filippi S, Fenton FH. 2013 Effects of pacing site and stimulation history on alternans dynamics and the development of complex spatiotemporal patterns in cardiac tissue. *Front. Physiol.* **4**, 71. (doi:10.3389/fphys.2013.00071)
79. Cherry EM, Fenton FH. 2004 Suppression of alternans and conduction blocks despite steep APD restitution: electrotonic, memory, and conduction velocity restitution effects. *Am. J. Physiol.-Heart Circ. Physiol.* **286**, H2332–H2341. (doi:10.1152/ajpheart.00747.2003)
80. Loppini A, Gizzi A, Cherubini C, Cherry EM, Fenton FH, Filippi S. 2019 Spatiotemporal correlation uncovers characteristic lengths in cardiac tissue. *Phys. Rev. E* **100**, 020201. (doi:10.1103/PhysRevE.100.020201)
81. Quarteroni A, Lassila T, Rossi S, Ruiz-Baier R. 2017 Integrated heart—coupling multiscale and multiphysics models for the simulation of the cardiac function. *Comput. Methods Appl. Mech. Eng.* **314**, 345–407. (doi:10.1016/j.cma.2016.05.031)
82. Krishnamoorthi S, Sarkar M, Klug WS. 2013 Numerical quadrature and operator splitting in finite element methods for cardiac electrophysiology. *Int. J. Numer. Methods Biomed. Eng.* **29**, 1243–1266. (doi:10.1002/cnm.2573)
83. Virag N, Jacquemet V, Henriquez CS, Zozor S, Blanc O, Vesin JM, Pruvot E, Kappenberger L. 2002 Study of atrial arrhythmias in a computer model based on magnetic resonance images of human atria. *Chaos* **12**, 754–763. (doi:10.1063/1.1483935)