

Data-Driven Estimation of Cardiac Electrical Diffusivity from 12-Lead ECG Signals

Oliver Zettinig^{a,b}, Tommaso Mansi^{a,*}, Dominik Neumann^{a,c}, Bogdan Georgescu^a, Saikiran Rapaka^a, Philipp Seegerer^{a,c}, Elham Kayvanpour^d, Farbod Sedaghat-Hamedani^d, Ali Amr^d, Jan Haas^d, Henning Steen^d, Hugo Katus^d, Benjamin Meder^d, Nassir Navab^b, Ali Kamen^a, Dorin Comaniciu^a

^a*Siemens Corporate Technology, Imaging and Computer Vision, Princeton, NJ, USA*

^b*Computer Aided Medical Procedures, Technische Universität München, Germany*

^c*Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany*

^d*Heidelberg University Hospital, Heidelberg, Germany*

Abstract

Diagnosis and treatment of dilated cardiomyopathy (DCM) is challenging due to a large variety of causes and disease stages. Computational models of cardiac electrophysiology (EP) can be used to improve the assessment and prognosis of DCM, plan therapies and predict their outcome, but require personalization. In this work, we present a data-driven approach to estimate the electrical diffusivity parameter of an EP model from standard 12-lead electrocardiograms (ECG). An efficient forward model based on a mono-domain, phenomenological Lattice-Boltzmann model of cardiac EP, and a boundary element-based mapping of potentials to the body surface is employed. The electrical diffusivity of myocardium, left ventricle and right ventricle endocardium is then estimated using polynomial regression which takes as input the QRS duration and electrical axis. After validating the forward model, we computed 9,500 EP simulations on 19 different DCM patients in just under three seconds each to learn the regression model. Using this database, we quantify the intrinsic uncertainty of electrical diffusion for given ECG features and show in a leave-one-patient-out cross-validation that the regression method is able to predict myocardium diffusion within the uncertainty range. Finally, our approach is tested on the 19 cases using their clinical

*tommaso.mansi@siemens.com – Tel: +1 609 734 3537 – Fax: +1 609 734 6565 – 755 College Road East, 08540 Princeton, NJ, USA

ECG. 84% of them could be personalized using our method, yielding mean prediction errors of 18.7 ms for the QRS duration and 6.5° for the electrical axis, both values being within clinical acceptability. By providing an estimate of diffusion parameters from readily available clinical data, our data-driven approach could therefore constitute a first calibration step toward a more complete personalization of cardiac EP.

Keywords: Cardiac Electrophysiology, Statistical Learning, Lattice-Boltzmann Method, Uncertainty Quantification, Electrocardiogram

1. Introduction

1.1. Clinical Rationale

With around 17.3 million deaths per year (Mendis et al., 2011), the global burden of cardiovascular diseases remains high and causes a significant social and economic impact. According to recent estimates, about 2% of adults in Europe (McMurray et al., 2012) and 2.4% of adults in the US (Roger et al., 2012) suffer from heart failure alone, with the prevalence rising to more than 10% among persons 70 years of age or older. One of the most common causes of heart failure is dilated cardiomyopathy (DCM), a condition with weakened and enlarged ventricles and atria, leading to an ineffective pump function that can directly and indirectly affect the lungs, liver, and other organ systems. The prevalence of DCM amounts to around 0.9% of adults in the US (Ferri, 2013), and the disease is the leading indication for heart transplantation in younger adults. Due to a large variety of individual causes and disease stages, diagnosis and treatment of DCM remains an open challenge.

Cardiac arrhythmia, i.e. irregular electrical activity of the heart, occurs frequently in heart failure patients, particularly in those with DCM (McMurray et al., 2012). But also beyond DCM, the prevalence of cardiac rhythm disorders has increased significantly in the last decade following an improvement in patient care (Marcus et al., 2013). Depending on the kind of rhythm disorder, which is commonly diagnosed using electrocardiography (ECG), the treatment of arrhythmia includes drug therapies, radio frequency ablation and the implantation of artificial pacemakers and cardioverter-defibrillators. Unfortunately, around 30% of patients are non-responders to these invasive treatments, and in up to 50% of the cases, recurrences are identified (Auricchio et al., 2011).

27 As a result, tools for a more predictive assessment of cardiac electrophys-
28 iology (EP) are needed. Computational assistance is not only required for
29 a superior patient management and diagnosis but could also benefit therapy
30 planning, outcome prediction and intervention guidance. While improved
31 risk stratification could help avoiding unnecessary surgeries, the potential of
32 optimizing invasive procedures, for instance by choosing optimal electrode
33 locations, can potentially lead to an increased success rate and fewer non-
34 responders. For this purpose, computational models can be employed to
35 study and evaluate patient-specific electrophysiology in-silico.

36 *1.2. Technical Background: Computational Models of Cardiac Electrophysi-* 37 *ology*

38 *1.2.1. Models of Cardiac Action Potential*

39 A wide range of computational models of cardiac EP with different bio-
40 logical scales and theoretical complexity has been proposed since the seminal
41 work of Hodgkin and Huxley (1952). Especially in the last decade, the com-
42 munity has witnessed tremendous progress in modeling efforts (Clayton et al.,
43 2011). Depending on their level of detail, EP models can be classified into
44 three groups: Biophysical, phenomenological and Eikonal models.

45 Biophysical cellular models capture cardiac electrophysiology directly at
46 cell level by describing biological phenomena responsible for myocyte de-
47 polarization and repolarization. More precisely, ionic interactions within the
48 cell and across the cell membrane (ion channels) are considered (Noble, 1962;
49 Luo and Rudy, 1991; Noble et al., 1998; Ten Tusscher et al., 2004) and lead
50 to complex equations, commonly one per molecular process. Although it has
51 been shown that biophysical models can reproduce different electrophysiol-
52 ogical behaviors such as action potential restitution and conduction velocity,
53 the large amount of parameters limits their usage in clinical applications due
54 to the difficulty of personalization.

55 Cell models are then integrated at the organ level using reaction-diffusion
56 partial differential equations (PDEs). Two major categories can be distin-
57 guished. While mono-domain approaches neglect interstitial effects and con-
58 sider the myocardium as single excitable tissue (Coudière and Pierre, 2006),
59 bi-domain strategies superimpose intra- and extra-cellular domains and take
60 different electrical properties into account (Bourgault et al., 2009). In the
61 absence of external stimuli, mono-domain models have been shown to pro-
62 duce almost identical results as their bi-domain counterparts (Potse et al.,
63 2006).

64 Phenomonological models, historically the first models to be proposed by
65 FitzHugh (1961), work at a more macroscopic level. Derived from experi-
66 mental observations, the action potential is described by a small number of
67 parameters with direct influence on its shape, disregarding the underlying
68 ionic interactions (Aliev and Panfilov, 1996; Mitchell and Schaeffer, 2003).
69 Having only few parameters with direct effect on measurable output facil-
70 itates model personalization, and the lower computational cost when com-
71 pared to biophysical models offers a reasonable compromise between model-
72 ing capacity and performance. The distinction between mono-domain organ
73 level integration schemes such as in Aliev and Panfilov (1996); Fenton and
74 Karma (1998); Mitchell and Schaeffer (2003) and bi-domain approaches such
75 as in Clayton and Panfilov (2008) can be applied to phenomonological mod-
76 els, too. Recent numerical advances based on Lattice-Boltzmann methods
77 (Rapaka et al., 2012) or Finite Element methods (Talbot et al., 2013) exploit
78 the massively parallel architecture of modern graphics processing units, and
79 allow near real-time performance and user interaction.

80 Eikonal models (Franzone et al., 1990; Keener and Sneyd, 1998; Serme-
81 sant et al., 2007) solely concentrate on the propagation of the electrical wave
82 to stimulate muscle activation. The formation as well as the shape of the
83 action potential in myocytes is neglected. Governed only by the anisotropic
84 speed of wave propagation, the local time of wave arrival throughout the
85 myocardium, can be computed very efficiently using fast marching methods
86 (Sethian, 1999; Wallman et al., 2012). While it has become possible to sim-
87 ulate wave reentry phenomena with Eikonal models (Pernod et al., 2011),
88 capturing other complex pathological conditions such as arrhythmias, fibril-
89 lations or tachycardia is more challenging.

90 1.2.2. Model Personalization

91 In order to apply the aforementioned EP models in clinical settings,
92 patient-specific physiology has to be captured by personalized model param-
93 eters. Finding those is challenging in the clinical workflow as the estimation
94 from patient data implies solving an inverse problem. In this context, the
95 *forward model* denotes the computation of the electrical wave propagation
96 from the heart to the point of measurement (catheter electrode, body sur-
97 face), and the *inverse model* the back-projection of measurement data onto
98 the heart and the inference of model parameters (Gulrajani, 1998).

99 Inverse problem techniques are computationally demanding because they
100 comprise an optimization problem and therefore require a large quantity of

101 forward model runs (Modre et al., 2002; Chinchapatnam et al., 2008; Dössel
102 et al., 2011). Alternatively, data-driven algorithms have been investigated
103 to tackle model personalization. Linking activation patterns with the result-
104 ing cardiac motion that can be observed in clinical images, Prakosa et al.
105 (2013) train a machine-learning algorithm to estimate depolarization times
106 for cardiac segments from regional kinematic descriptors. Jiang et al. (2011)
107 apply statistical learning to map body surface potentials onto the epicardium.
108 Konukoglu et al. (2011) derive a surrogate EP model based on polynomial
109 chaos theory to personalize an Eikonal model. Wallman et al. (2014) infer
110 tissue conduction properties using Bayesian inference to be patient-specific.
111 The advantage of these statistical methods is the possibility to quantify un-
112 certainty and to optimize the location of measurements. Machine learning
113 techniques could therefore constitute efficient strategies for model personal-
114 ization. However, a sufficient sampling of the parameter space is needed for
115 these approaches to yield meaningful results. In this study, we aim to achieve
116 an estimation of model parameters only from sparse electrocardiogram data.

117 *1.2.3. Models of Electrocardiogram and Torso Potential*

118 From the perspective of data acquisition, endocardial mapping (Sermesant et al., 2009; Relan et al., 2011) facilitates the parameter estimation as it
119 provides dense potential measurements but it is pre-operatively often avoided
120 as it is invasive. A non-invasive alternative is to back-project electrical poten-
121 tials measured at the body surface in the form of electrocardiograms (ECG),
122 to the epicardium. Considering the ill-posedness of the parameter estimation,
123 the use of body surface mapping (BSM) has been investigated (Dössel et al.,
124 2011; Wang et al., 2011; Han et al., 2013). In contrast to standard 12-lead
125 ECG, BSM is however not yet widely available as diagnostic modality.
126

127 If body surface ECG data is used for parameter estimation, regardless
128 of the number of traces, a model of electrical potentials at the surface of
129 the torso is needed. In terms of the forward model, current approaches
130 employ both Finite Element (FEM) and Boundary Element (BEM) methods.
131 While the former intrinsically allow varying conductivity within and across
132 different organs (Li et al., 2007; Geneser et al., 2008; Liu et al., 2012), the
133 latter either assume constant isotropic conductivity throughout the entire
134 torso (Barr et al., 1977; Shou et al., 2009) or integrate additional surface
135 meshes delineating neighboring organs (Potse et al., 2009). Furthermore, in
136 a numerical study by Boulakia et al. (2010), decoupling the computation
137 of cardiac electrophysiology and body surface potentials has been shown to

138 preserve the shape of ECG features well.

139 *1.3. Aim of the Study*

140 The personalization of computational EP models, i.e. the estimation
141 of patient-specific model parameters, remains challenging because of lack
142 of dense data and the ill-posedness of the inverse problem. We therefore
143 propose in this work to estimate EP model parameters from standard 12-lead
144 electrocardiograms (ECG) only using a data-driven method that provides
145 insight into estimation uncertainty. In particular, our method is based on
146 statistical learning and employs polynomial regression to map ECG features
147 to model parameters, instead of finding a solution of the inverse problem
148 numerically. The key contributions of this work are:

- 149 i A fast forward model of cardiac electrophysiology and electrocardiogram
150 based on a Lattice-Boltzmann formulation and the boundary element
151 method.
- 152 ii A novel data-driven approach to automatically and efficiently estimate
153 heart electrical diffusivity from 12-lead ECG features.
- 154 iii The quantification of the intrinsic uncertainty of the inverse problem,
155 i.e. the uncertainty of myocardial diffusion given a set of ECG features,
156 through statistical learning.
- 157 iv The evaluation of our estimation framework on 19 DCM cases.

158 This study extends our previous work (Zettinig et al., 2013a) as follows:

- 159 i In a detailed quantitative evaluation and convergence analysis of the em-
160 ployed forward model and its parts, we show the influence of various
161 model parameters on the ECG features.
- 162 ii Results of benchmark experiments allow insights on the computational
163 performance of our approach.
- 164 iii A comparison with other statistical learning techniques justifies the choice
165 of multivariate polynomial regression.
- 166 iv For a quantitative evaluation of the diffusion estimation method, we use
167 a significantly more extensive dataset of synthetic and real case data,
168 allowing to capture a bigger variety of individual physiologies.

169 **2. Methods**

170 This section presents the details of the proposed data-driven EP pa-
171 rameter estimation framework. Section 2.1 describes how a patient-specific

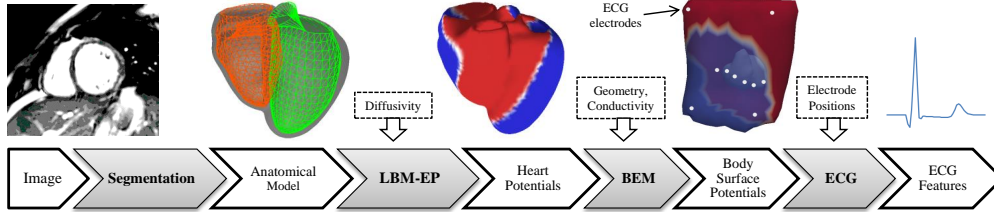


Figure 1: Steps of proposed forward model of ECG.

172 anatomical model is derived from clinical images. In Sec. 2.2, a fast forward
 173 model of cardiac electrophysiology, body surface potentials and electrocar-
 174 diogram, as shown in Fig. 1, is detailed. Thereafter, Sec. 2.3 describes the
 175 proposed data-driven diffusion estimation procedure. Implementation details
 176 are reported in Sec. 2.4.

177 2.1. Patient-Specific Model of Cardiac Anatomy

178 The complete workflow of anatomical model generation is depicted in
 179 Fig. 2. First, we employ the framework presented in Zheng et al. (2008)
 180 to automatically estimate, under expert guidance, heart morphology from
 181 cine magnetic resonance images (MRI). For anatomical structure localiza-
 182 tion, the Marginal Space Learning (MSL) framework intuitively finds control
 183 points representing important landmarks such as valves and ventricular sep-
 184 tum cusps using Haar- and steerable features. Then, a point-distribution
 185 model of biventricular geometry is mapped to these control points and suc-
 186 cessively deformed according to learning-based boundary delineation through
 187 a Probabilistic Boosting Tree (PBT). Using a manifold-based motion model,
 188 the resulting surface meshes are tracked over the cardiac sequence such that
 189 point correspondences are maintained. To form a closed surface mesh of the
 190 biventricular myocardium, the segmented triangulations of the epicardium
 191 and endocardia are fused together. The myocardium at end-diastole is fi-
 192 nally mapped onto a Cartesian grid with isotropic spacing and represented
 193 as a level-set.

194 Based on the original segmentation meshes and point-to-point distances,
 195 we consider five domains in our anatomical model: The left and right ventric-
 196 ular septum, which mimics the His bundle and serves as initialization zone
 197 of the electrophysiological wave, the left and right endocardia mimicking the

198 Purkinje system of fast electrical diffusivity, and finally the myocardium with
 199 slower diffusivity.

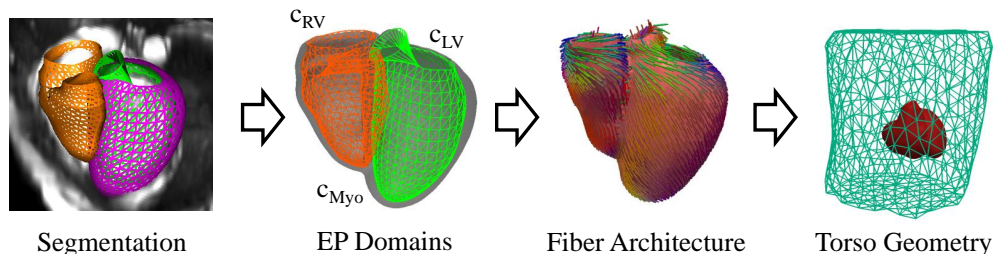


Figure 2: Workflow of anatomical model generation. *See text for details.*

200 As diffusion tensor imaging (DTI) is not yet clinically available (Wu et al.,
 201 2009), the rule-based strategy described by Bayer et al. (2012) is extended
 202 as proposed by Zettinig et al. (2013b) to compute a generic model of myo-
 203 cardium fiber architecture. Below the basal plane, identified automatically
 204 using the point correspondences of the initial triangulations, the fiber eleva-
 205 tion angle α_f is assigned to all grid nodes. Defined as the angle with respect
 206 to the short axis, α_f varies linearly across the myocardium from -70° on the
 207 epicardium to $+70^\circ$ on the endocardium. Around the valves, fiber directions
 208 are fixed (longitudinal around the aortic valve, tangential otherwise), and
 209 between the basal plane and the valves finally interpolated first following
 210 the myocardium surface, then transmurally (Moireau, 2008; Zettinig et al.,
 211 2013b). All interpolations throughout the myocardium rely on geodesic dis-
 212 tances and the Log-Euclidean framework (Arsigny et al., 2006). Figure 2
 213 illustrates the myocardium fiber model and the electrophysiology zones.

214 A body surface triangulation is obtained using a manual, two-step pro-
 215 cedure. First, the contours of the torso are outlined in coronal, sagittal and
 216 transverse slices of the survey MR image, and visualized together with the
 217 heart model. Second, a manual affine registration of an atlas of torso geome-
 218 try to the contours is performed as illustrated in Fig. 3. The atlas is obtained
 219 from a full-body CT dataset of a subject within normal weight range.

220 2.2. Fast Forward Model of Cardiac Electrocardiogram

221 Our forward model consists of three sequential steps described in the fol-
 222 lowing sections (Fig. 1). First, we compute cardiac electrophysiology using
 223 the LBM-EP algorithm proposed by Rapaka et al. (2012). Second, we esti-
 224 mate extracellular potentials at the epicardium using an elliptic formulation

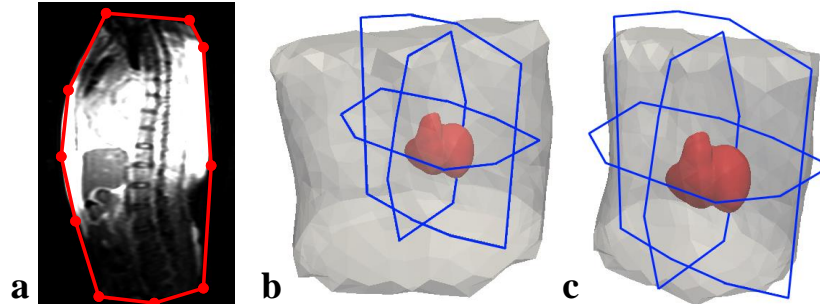


Figure 3: **a)** Sagittal image slice and manually outlined contour. **b)** Atlas of torso geometry before registration and **c)** after manual registration to body contours in coronal, sagittal and transverse image slices.

225 and project them to the torso by means of a Boundary Element Method
 226 technique. Ultimately, ECG traces are computed and ECG features auto-
 227 matically calculated.

228 *2.2.1. LBM-EP: Lattice-Boltzmann Model of Myocardium Transmembrane*
 229 *Potentials*

230 Cardiac EP is computed according to the phenomenological mono-domain
 231 model proposed by Mitchell and Schaeffer (2003), which describes the nor-
 232 malized transmembrane potential (TMP) $v(t) \in [0, 1]$ throughout the my-
 233 ocardium with the following equation:

$$\frac{\partial v}{\partial t} = J_{in} + J_{out} + J_{stim} + c\nabla \cdot D\nabla v \quad (1)$$

234 Electrical diffusion is formulated anisotropically with the diffusion coef-
 235 ficient c and the anisotropy ratio ρ , defining the anisotropic diffusion tensor
 236 $D = \rho I + (1 - \rho)\mathbf{f}\mathbf{f}^T$ with \mathbf{f} denoting the fiber direction. The EP zones as
 237 defined in Sec. 2.1 are assigned three different diffusion coefficients: c_{LV} and
 238 c_{RV} for the left and right endocardium, respectively (fast conducting Purkinje
 239 network), and c_{Myo} for the myocardium.

The model simplifies all ion channel interactions to only an inward current J_{in} and an outward current J_{out} (Eqs. 2-3). The former captures the fast acting ionic currents in the myocyte and depends on the gating variable $h(t)$

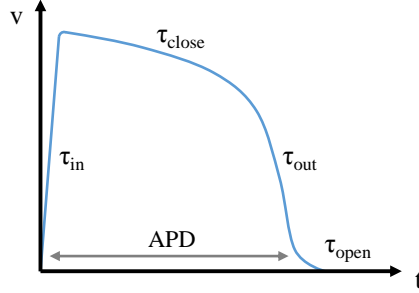


Figure 4: Four stages of the myocyte action potential and the relating parameters of the Mitchell-Schaeffer model.

that models the state of the ion channels.

$$J_{in} = \frac{h(t)v^2(1-v)}{\tau_{in}}, \text{ with } \frac{dh}{dt} = \begin{cases} \frac{1-h}{\tau_{open}}, & \text{if } v < v_{gate} \\ \frac{-h}{\tau_{close}}, & \text{otherwise} \end{cases} \quad (2)$$

$$J_{out} = \frac{-v}{\tau_{out}} \quad (3)$$

240 The time constants $\tau_{in} \ll \tau_{out} \ll \tau_{open}, \tau_{close}$ are directly related to the
 241 shape and duration of the action potential, allowing for personalization from
 242 clinical data. As illustrated in Fig. 4, τ_{close} relates to the action potential du-
 243 ration (APD), for which a linear transmural gradient as described by Glukhov
 244 et al. (2010) is employed. The remaining model parameters, including the
 245 change-over voltage v_{gate} , are obtained from literature (Mitchell and Schaeffer,
 246 2003) and kept constant throughout the myocardium. Table 1 lists all
 247 fixed model parameters.

Table 1: Parameters used for the Mitchell-Schaeffer model (Mitchell and Schaeffer, 2003; Glukhov et al., 2010). Note that v_{gate} is dimensionless because $v(t)$ is normalized to $[0, 1]$.

Parameter	
v_{gate}	0.13
τ_{in}	0.3 ms
τ_{out}	6 ms
τ_{open}	120 ms
$\tau_{close_{endo}}$	130 ms
$\tau_{close_{epi}}$	90 ms

The complex PDE (Eq. 1) is solved using the LBM-EP algorithm, an efficient Lattice-Boltzmann method, proposed by Rapaka et al. (2012). It should be noted, though, that the LBM-EP algorithm is generic and would allow any mono-domain cell model to be solved. In short, the method maintains a vector of distribution functions $\mathbf{f}(\mathbf{x}) = \{f_i(\mathbf{x})\}_{i=1..7}$, where $f_i(\mathbf{x})$ represents the probability of finding a particle traveling along the edge \mathbf{e}_i of node \mathbf{x} . The seven indices correspond to the central position and the six principal connections on the Cartesian grid, respectively. Its computation is decomposed into two consecutive steps, namely the *collision* phase, yielding intermediate post-collision states f_i^* and the *streaming* phase, propagating the distribution functions along their corresponding edges:

$$f_i^* = f_i - A_{ij}(f_j - \omega_j v) + \delta t \omega_i (J_{in} + J_{out} + J_{stim}), \quad (4)$$

$$f_i(\mathbf{x} + \mathbf{e}_i, t + \delta t) = f_i^*(\mathbf{x}, t) \quad (5)$$

248 The collision matrix $A = (A_{ij})$ relaxes the distribution function f_i toward
 249 the local value of the potential v and is defined such that anisotropic fiber-
 250 related diffusion is taken into account. The weighting factors ω_i are utilized
 251 to emphasize the center position. We refer the reader to Rapaka et al. (2012)
 252 for further details. Using a forward Euler scheme, the gating variable $h(t)$ can
 253 easily be updated at every node. Eventually, the transmembrane potential
 254 $v(\mathbf{x}, t)$ is defined as the sum of the distribution functions: $v(\mathbf{x}, t) = \sum_i f_i(\mathbf{x}, t)$
 255 and transferred to the range $[-70 \text{ mV}, 30 \text{ mV}]$ using the scaling factors given
 256 in Mitchell and Schaeffer (2003). The depolarization times $T_d(\mathbf{x})$ are obtained
 257 as the points in time when the potential first exceeds the change-over voltage:

$$T_d(\mathbf{x}) = \arg \min_t \{v(\mathbf{x}, t) \geq v_{gate}\} \quad (6)$$

258 2.2.2. Boundary Element Model of Torso Potentials

259 For the propagation of electrical potentials through the body, it is neces-
 260 sary to estimate cardiac extracellular potentials $\phi_e(t)$ from the TMP $v(t)$. To
 261 that end, we employ the elliptic formulation proposed by Chhay et al. (2012),
 262 which assumes a constant diffusion anisotropy ratio $\lambda = c_i(\mathbf{x})/c_e(\mathbf{x})$, with c_i
 263 and c_e denoting the intra- and extracellular diffusion coefficients respectively.
 264 Within the entire myocardium domain Ω , the extracellular potential ϕ_e is ex-
 265 pressed as:

$$\phi_e(\mathbf{x}, t) = \frac{\lambda}{1 + \lambda} \frac{1}{|\Omega|} \int_{\Omega} (v(\mathbf{y}, t) - v(\mathbf{x}, t)) d\mathbf{y} \quad (7)$$

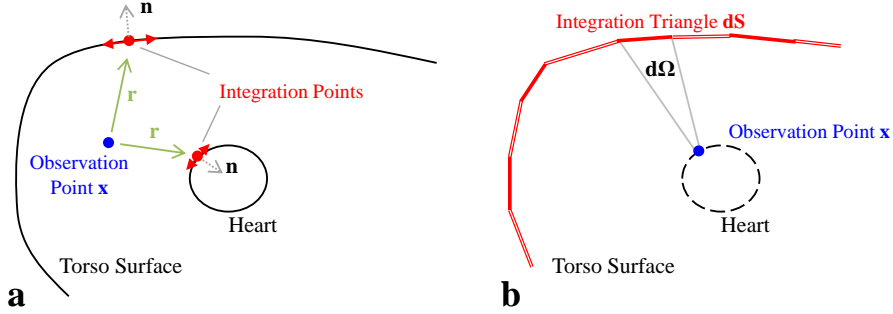


Figure 5: **a)** For any given observation point in the thoracic domain, both torso and heart surfaces need to be integrated as defined in Eq. 8. **b)** After discretization, geometric coefficients of the P matrices in Eq. 10 require the evaluation of solid angles.

266 Next, we utilize a boundary element method (BEM) as described in Barr
 267 et al. (1977) and refined in Shou et al. (2009) to project the potentials ϕ_e from
 268 the epicardium to the torso. Before, tri-linear interpolation is used to map
 269 ϕ_e from the Cartesian grid back to the epicardial surface mesh. Following
 270 Green's second identity, the potential $\phi(\mathbf{x})$ at any observation point \mathbf{x} of the
 271 thoracic domain is given as:

$$\phi(\mathbf{x}) = \frac{1}{4\pi} \int_{S_B} \phi_B \frac{\mathbf{r} \cdot \mathbf{n}}{\|\mathbf{r}\|^3} dS_B + \frac{1}{4\pi} \int_{S_H} \left[\phi_e \frac{\mathbf{r} \cdot \mathbf{n}}{\|\mathbf{r}\|^3} + \frac{\nabla \phi_e \cdot \mathbf{n}}{\|\mathbf{r}\|} \right] dS_H \quad (8)$$

272 Hereby, subscripts B denote the body surface and the potentials there-
 273 upon, S_H the epicardial heart surface. The surface normals \mathbf{n} face outward
 274 of the domain under consideration (i.e. outward at the torso and inward at
 275 the epicardium). \mathbf{r} is defined as the vector from \mathbf{x} to the point of integration
 276 as illustrated in Fig. 5a. Note that Eq. 8 assumes that $\nabla \phi_B = 0$.

After placing the observation point \mathbf{x} only onto the two surfaces, discretization in triangular meshes, and reformulation in matrix form, a system of linear equations can be constructed (Barr et al., 1977):

$$\mathbf{P}_{BB} \phi_B + \mathbf{P}_{BH} \phi_e + \mathbf{G}_{BH} \Gamma_H = 0 \quad (9)$$

$$\mathbf{P}_{HB} \phi_B + \mathbf{P}_{HH} \phi_e + \mathbf{G}_{HH} \Gamma_H = 0 \quad (10)$$

277 Obtaining the geometric coefficients of matrices P and G requires the
 278 evaluation of two integrals. The integral $\int (\mathbf{r} \cdot \mathbf{n}) / \|\mathbf{r}\|^3 dS$ in fact describes
 279 the solid angle $d\Omega$ subtended at any observation point by a surface element

280 dS (see Fig. 5b), and can be efficiently computed with the following closed
 281 form formula (Van Oosterom and Strackee, 1983):

$$\tan \frac{d\Omega}{2} = \frac{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}{\prod_{i=1}^3 \|\mathbf{a}_i\| + \|\mathbf{a}_1\| (\mathbf{a}_2 \cdot \mathbf{a}_3) + \|\mathbf{a}_2\| (\mathbf{a}_3 \cdot \mathbf{a}_1) + \|\mathbf{a}_3\| (\mathbf{a}_1 \cdot \mathbf{a}_2)} \quad (11)$$

282 Hereby, vectors \mathbf{a}_i denote the vectors from the observation point to the three
 283 vertices of the triangulated surface element dS . The surface-over-distance in-
 284 tegral $\int dS/\mathbf{r}$, on the other hand, is solved using Gaussian quadrature. Can-
 285 celing out the matrix Γ_H , which contains the gradients $\nabla\phi_e$, a precomputable
 286 transfer matrix that entirely depends on the geometry can be defined:

$$\mathbf{Z}_{BH} = (\mathbf{P}_{BB} - \mathbf{G}_{BH}\mathbf{G}_{HH}^{-1}\mathbf{P}_{HB})^{-1} (\mathbf{G}_{BH}\mathbf{G}_{HH}^{-1}\mathbf{P}_{HH} - \mathbf{P}_{BH}) \quad (12)$$

287 This allows to express body surface potentials by means of a simple matrix
 288 multiplication: $\phi_B = \mathbf{Z}_{BH} \phi_e$.

289 2.2.3. Electrocardiogram Calculation

290 From the potentials ϕ_B at the torso, the standard Einthoven, Goldberger
 291 and Wilson leads (Chung, 1989) are computed. For the sake of simplicity,
 292 electrode positions were chosen to coincide with manually selected torso mesh
 293 vertex positions.

294 In this work, we focus on two meaningful ECG features. On the one hand,
 295 the duration of the QRS complex Δ_{QRS} is dependent on the total time the
 296 electrical wave requires to propagate throughout the entire myocardium. On
 297 the other hand, the mean electrical axis angle α is suited to detect imbalances
 298 between left and right ventricular wave conduction. From the computed ECG
 299 signals, Δ_{QRS} and α are derived as follows:

- 300 • For numerical stability, the QRS complex is detected using the depo-
 301 larization times computed by LBM-EP. Assuming one full heart cycle,
 302 it is computed as follows: $\Delta_{QRS} = \max_x T_d(\mathbf{x}) - \min_x T_d(\mathbf{x})$.
- 303 • The electrical axis is computed using the Einthoven leads I and II:
 304 $\alpha = \arctan [(2h_{II} - h_I)/(\sqrt{3}h_I)]$, where the h_i 's are the sum of the
 305 automatically detected R and S peak amplitudes (minimum and max-
 306 imum) in the respective leads during the QRS complex.

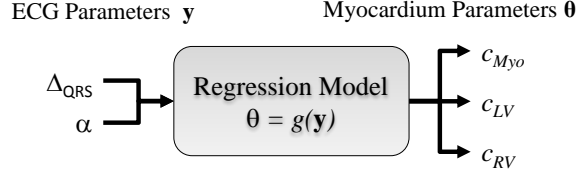


Figure 6: Schematic diagram of the data-driven backward ECG model.

307 *2.3. Data-Driven Estimation of Myocardium EP Diffusion*

308 The forward model as described above can be seen as a dynamic system
 309 $\mathbf{y} = f(\theta)$ with the diffusion coefficients $\theta = (c_{Myo}, c_{LV}, c_{RV})$ as free param-
 310 eters and the ECG features $\mathbf{y} = (\Delta_{QRS}, \alpha)$ as outputs of the system. Estimati-
 311 ing diffusion parameters from ECG features therefore consists in evaluating
 312 a function $g(\mathbf{y})$ that approximates the inverse problem $\theta = g(\mathbf{y}) \approx f^{-1}(\mathbf{y})$,
 313 as shown in Fig. 6. In contrast to solving the inverse problem numerically
 314 using an optimization strategy such as Dössel et al. (2011), we propose to
 315 learn the inverse function instead.

Table 2: Diffusion coefficient configurations for normalization forward runs.

Configuration	Diffusion coefficients (mm ² /s)		
	c_{Myo}	c_{LV}	c_{RV}
F_1	100	4,900	4,900
F_2	100	100	4,900
F_3	100	4,900	100

316 The ECG features Δ_{QRS} and α vary significantly within the population,
 317 even in healthy subjects, due to a variety of factors including heart mor-
 318 phology and position. To cope with this geometrical variety, our algorithm
 319 scouts the parameter space using three forward model runs with the pre-
 320 defined diffusion coefficients listed in Tab. 2. The resulting ECG features
 321 are then used for an effective normalization scheme, intrinsically considering
 322 geometrical features of a particular patient:

- 323 • Configuration F_1 contains nominal EP diffusion parameters and thus
 324 entails a normal wave propagation. Provided the same diffusivity, the
 325 electrical wave will take longer to propagate through the entire my-
 326 ocardium in larger hearts, which is why we use $\Delta_{QRS_{F_1}}$ to normalize
 327 the QRS duration: $\overline{\Delta_{QRS}} = \Delta_{QRS} / \Delta_{QRS_{F_1}}$.

328 • The other two configurations contain extremely low LV and RV diffu-
 329 sivity (LBBB-like scenario: F_2 ; RBBB-like scenario: F_3). The obtained
 330 electrical axis parameters α_{F_2} and α_{F_3} scout the patient-specific space
 331 of axis deviation, because we assume that the vast majority of forward
 332 model runs with arbitrary physiological diffusion coefficients will yield
 333 an electrical axis between them: $\alpha_{F_2} \leq \alpha \leq \alpha_{F_3}$. Therefore, we perform
 334 normalization as follows: $\bar{\alpha} = (\alpha - \alpha_{F_2}) / (\alpha_{F_3} - \alpha_{F_2})$.

335 Note that our normalization scheme does not lead to the same numerical
 336 ranges of $\overline{\Delta_{QRS}}$ and $\bar{\alpha}$. The purpose of the normalization is only to com-
 337 pensate for inter-patient variability; the regression framework will cope with
 338 scaling of the input values. Finally, multivariate polynomial regression of
 339 degree N is employed to learn the model $\theta = g(\overline{\Delta_{QRS}}, \bar{\alpha})$. One regression
 340 function of the form

$$g(\overline{\Delta_{QRS}}, \bar{\alpha}) = \sum_{i=0}^N \sum_{j=0}^N \beta_{i,j} (\overline{\Delta_{QRS}})^i (\bar{\alpha})^j + \varepsilon \quad (13)$$

341 is learned for each diffusivity parameter independently, $\mathbf{g} = (g_{Myo}, g_{LV}, g_{RV})$.
 342 During training, the regression coefficients $\beta_{i,j}$ are determined using QR de-
 343 composition such that the data is explained with minimal error ε . During
 344 testing, the diffusivity parameters are estimated for unseen data using mea-
 345 sured and normalized ECG features:

$$\begin{pmatrix} c_{\hat{M}yo} \\ c_{\hat{L}V} \\ c_{\hat{R}V} \end{pmatrix} = \begin{bmatrix} g_{Myo} \\ g_{LV} \\ g_{RV} \end{bmatrix} \begin{pmatrix} \overline{\Delta_{QRS}} \\ \bar{\alpha} \end{pmatrix} \quad (14)$$

346 2.4. Implementation

347 The strictly local stream-and-collide rules of the LBM-EP algorithm are
 348 inherently node-wise and can be implemented very efficiently in a single ker-
 349 nel on a GPU architecture. We use NVIDIA CUDA¹, version 5.5, as our
 350 development environment. As shown by Georgescu et al. (2013), the sim-
 351 ulation of transmembrane and extracellular potentials for a complete heart
 352 cycle on a Cartesian grid with an isotropic resolution of 1.5 mm only requires

¹Compute Unified Device Architecture, <http://developer.nvidia.com/cuda-toolkit>

353 ≈ 3 seconds on an NVIDIA GeForce GTX 580 graphics card. The boundary
354 element solver relies on the C++ Eigen library (Guennebaud et al., 2010).
355 Training of and prediction with the regression model was performed using
356 the MATLAB and Statistics Toolbox Release 2013b (MathWorks, Inc.)

357 **3. Experiments and Results**

358 Before evaluating the method, clinical acceptance criteria were defined. In
359 a study by Surawicz et al. (2009), the normal range for QRS duration in adult
360 males was found to be between 74 and 114 ms (average 95 ms). Investigating
361 intra-patient variability in electrocardiograms, Michaels and Cadoret (1967)
362 defined the maximum permissible day-to-day difference in QRS duration as
363 20 ms. We assume predictions of QRS durations to be successful if within this
364 range. The electrical axis, which is dependent on age and body physiology, is
365 considered to be normal within -30° and 90° . A rough rule-based diagnosis
366 scheme is often applied in clinical routine. Left axis deviation, for instance,
367 is present if lead I is positive and aVF is negative. Because such a scheme
368 divides the QRS front plane in sectors of no less than 30° , prediction of the
369 electrical axis was assumed to be successful if within this range.

370 *3.1. Evaluation of the Proposed Forward Model*

371 A quantitative evaluation of the proposed forward model was carried out
372 to understand model behavior but also to identify the optimal numerical
373 parameters. For an extensive analysis of the LBM-EP solver, the reader is
374 referred to Rapaka et al. (2012). The following sections therefore focus on
375 the mapping of cardiac potentials onto the body surface and on the impact
376 of EP parameters on the computed ECG.

377 *3.1.1. Quantitative Evaluation and Convergence Analysis of Torso Mapping*

378 For the evaluation of the boundary element mapping from the epicardium
379 to the torso, we chose a setup where an analytical solution to Eq. 8 exists.
380 Both epicardium and torso were assumed to be concentric spheres, with radii
381 r_H for the heart sphere and r_B for the body sphere. If not stated otherwise,
382 Gaussian quadrature of order 37 was used. Homogeneous material between
383 the two surfaces was also assumed. Using a spherical coordinate system with
384 θ as the polar angle to the Cartesian z-direction and φ as the azimuth angle
385 in the x-y-plane (Fig. 7a), we defined the extracellular potentials on the

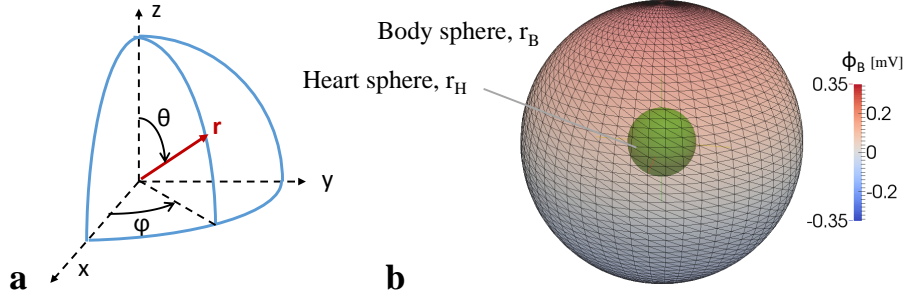


Figure 7: **a)** Definition of spherical coordinate system. **b)** Heart sphere (green) and body sphere (semi-transparent) with mapped potentials ϕ_B .

386 heart surface: $\phi_e(\theta, \varphi) = \cos(\theta)$ mV. Then, the potentials on the body ϕ_B
 387 were given as:

$$\phi_B(\theta, \varphi) = \frac{3}{2} \cos(\theta) \frac{1}{\frac{r_H}{r_B} + \frac{r_B^2}{2r_H^2}} \text{ mV} \quad (15)$$

388 Fig. 7b illustrates the mapped potentials on the body sphere. In the
 389 reported experiments, $r_H = 100$ mm to roughly represent the human heart.
 390 Figure 8a reports computed and analytical body potentials throughout a
 391 body sphere with $r_B = 300$ mm. Mapping to different body spheres (see
 392 Tab. 3 for mesh resolution details) showed that the algorithm was able to
 393 correctly compute the potentials at various distances (Fig. 8b). Absolute
 394 errors were on average $4.1 \cdot 10^{-5} \pm 1.4 \cdot 10^{-4}$ mV (mean \pm standard deviation),
 395 far below the clinical acceptance threshold.

Table 3: Body spheres used for torso mapping evaluation. See text for details.

Radius r_B (mm)	Number of Vertices	Avg. Edge Length l (mm)
150	3,482	10.5
300	3,482	21.0
400	3,482	28.0
600	3,482	42.0

396 A convergence analysis with respect to the mesh resolution (Fig. 8c) in-
 397 dicated that with around 1,500 mesh vertices (average edge length 31.7 mm
 398 for $r_B = 300$ mm) the ratio between BEM-based and analytical solutions
 399 is 99.70%, which relates to an absolute error of on average $9.3 \cdot 10^{-4}$ mV
 400 and is below the sensitivity of ECG sensors. For subsequent experiments,

401 we therefore choose a similar resolution of 30 to 35 mm for the torso mesh.
 402 Similarly, the relative error also converged with increasing order of Gaussian
 403 quadrature as expected (Fig. 8d, experiments conducted with highest resolution
 404 quadrature mesh). Yet, low orders already reached a high degree of precision. The
 405 following experiments were therefore carried out with an order of 6, which
 406 showed to be a good compromise between accuracy and runtime performance
 407 (more than $6\times$ as fast as highest order under consideration).

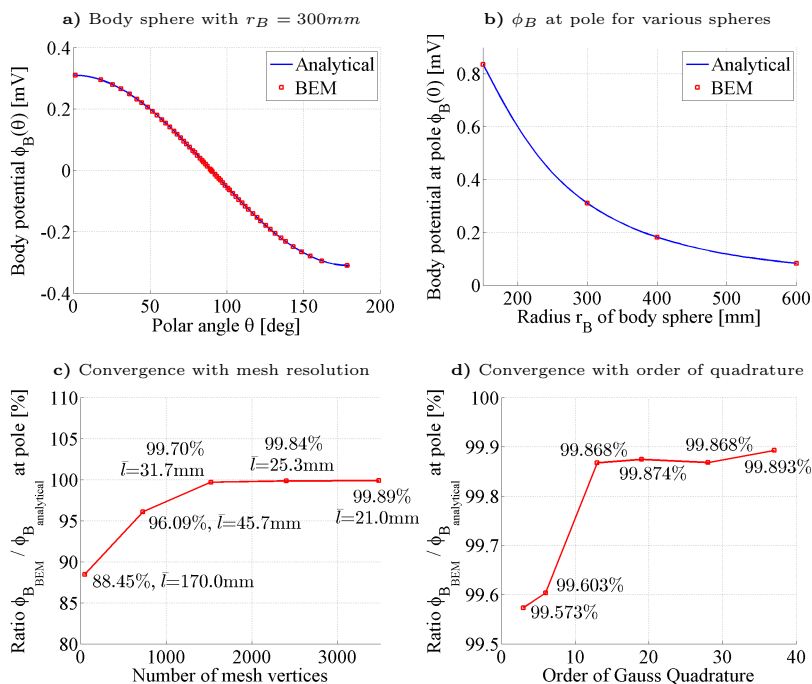


Figure 8: Evaluation of BEM torso mapping. **a)** Potentials throughout a body sphere with $r_B = 300$ mm and **b)** potentials at the pole ($\theta = 0$, location of maximum error) for various body spheres (Tab. 3) matched the analytical solution. **c)** Ratio between BEM-based and analytical solution for various mesh resolutions and average edge lengths \bar{l} , **d)** and for various orders of Gaussian quadrature, showing that the method converges rapidly with increasing mesh resolution and order of quadrature.

408 3.1.2. Parameter Evaluation of Complete Forward Model

409 Understanding the input parameters and output feature space of a given
 410 model is crucial before applying machine learning techniques and performing
 411 predictions. Therefore, we evaluated the influence of the most important

412 parameters of our forward model on the ECG features under consideration.
 413 On a representative patient case, the dependence of QRS duration Δ_{QRS} and
 414 electrical axis α on diffusivity c , action potential duration (APD; governed by
 415 τ_{close}) and fiber elevation angle $\angle f$ was studied. In the following experiments,
 416 each of these parameters was varied, while the other parameters were fixed
 417 to their nominal value as given in Tab. 1-2.

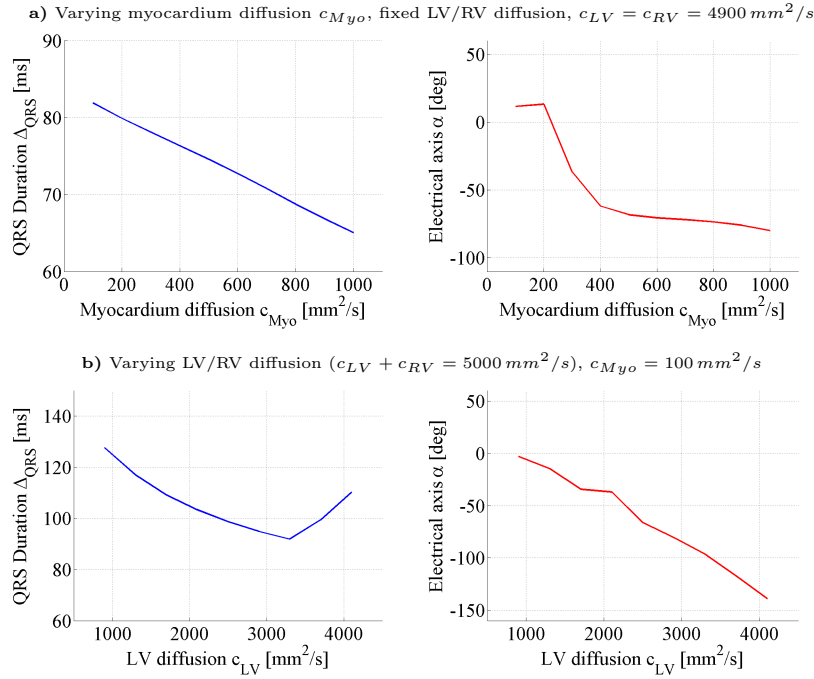


Figure 9: Influence of diffusivity coefficients on ECG features. **a)** QRS duration was linearly dependent on myocardium diffusion when LV and RV diffusion were fixed. Electrical axis varied little except for very low myocardium diffusion. **b)** Electrical axis was almost linearly dependent on LV/RV diffusion when myocardium diffusion was fixed. QRS duration varied little except for low LV or RV diffusion ($< 1,500 \text{ mm}^2/\text{s}$, borders of plot).

418 First, the forward model was run with myocardium diffusion c_{Myo} ranging
 419 from 100 to 1,000 mm^2/s , LV and RV diffusivity were fixed. As illustrated
 420 in Fig. 9a, and as expected, the QRS duration showed linear dependence on
 421 c_{Myo} . After an initial drop, the electrical axis stabilized when $c_{Myo} > 400$
 422 mm^2/s . In total, the range of α in this experiment accounted for 92.9°.
 423 This behavior was not surprising, because the electrical axis was entirely
 424 governed by LV and RV endocardium diffusion at very low c_{Myo} values. As

425 c_{Myo} increased, the effect of the Purkinje model was complemented by the
 426 fast myocardial diffusivity and the electrical axis value depended mostly on
 427 the geometrical configuration of the heart, i.e. its shape and relative position
 428 in the torso.

429 c_{Myo} was then fixed while the endocardial diffusion (c_{LV} , c_{RV}) was varied
 430 between 100 and 4,900 mm²/s with the constraint $c_{LV} + c_{RV} = 5,000$ mm²/s.
 431 As illustrated in Fig 9b, an almost linear dependence of the electrical axis on
 432 endocardial diffusion was observed. This was also expected, as diffusion dif-
 433 ferences in the Purkinje fibers intuitively change the depolarization pattern.
 434 When c_{LV} and c_{RV} were similar ($|c_{LV} - c_{RV}| < 1,000$ mm²/s), the depolar-
 435 ization was controlled by the Purkinje system, and a small range of 8.6 ms
 436 was observed for the QRS duration. Either c_{LV} or c_{RV} approaching a bundle
 437 branch block scenario and thus becoming closer to myocardial diffusion, the
 438 QRS duration increased. The total range of Δ_{QRS} was found to be 35.8 ms.

439 Next, the influence of different spatial distributions of τ_{close} parameters,
 440 which control action potential duration, was investigated. The employed lin-
 441 ear transmural gradient of action potential duration (Glukhov et al., 2010)
 442 was compared to a model with M-cells as described by Wilson et al. (2011).
 443 Hereby, we assumed M-cells to be located in the center of the myocardial
 444 wall, $\tau_{close_{mid}} = 110\% \cdot \tau_{close_{endo}}$ and performed linear interpolation between
 445 endocardium and M-cells, and between M-cells and epicardium to obtain a
 446 spatially varying map of τ_{close} values. For both scenarios (linear gradient
 447 throughout the wall and the M-cell model), we defined the parameter r_T
 448 as the ratio between the APD parameters at epicardium and endocardium:
 449 $r_T = \tau_{close_{epi}} / \tau_{close_{endo}}$. Figure 10a shows that the difference in the considered
 450 ECG features between the linear gradient model and the used M-cell model
 451 was marginal. This result was not surprising as τ_{close} controls cardiac repo-
 452 larization, whereas the QRS duration and electrical axis depend mostly on
 453 cardiac depolarization. Furthermore, we also analyzed how regional differ-
 454 ences in APD can influence the ECG parameters. To that end, we created a
 455 base-to-apex gradient by defining an additional ratio $r_{BA} = \tau_{close_{base}} / \tau_{close_{apex}}$.
 456 In Fig. 10b, the resulting variation in Δ_{QRS} and α is illustrated. In this case,
 457 QRS duration was, as expected, only minimally affected (range 0.15 ms) but
 458 the electrical axis varied by 13.6°.

459 Finally, we investigated the effect of the fiber model on the ECG features.
 460 According to the study by Lombaert et al. (2012), fiber angles in human phys-
 461 iology range on average from about 50° to 80° on the epicardium. As shown
 462 in Fig. 11, the variation of Δ_{QRS} in that range was small, with a range of

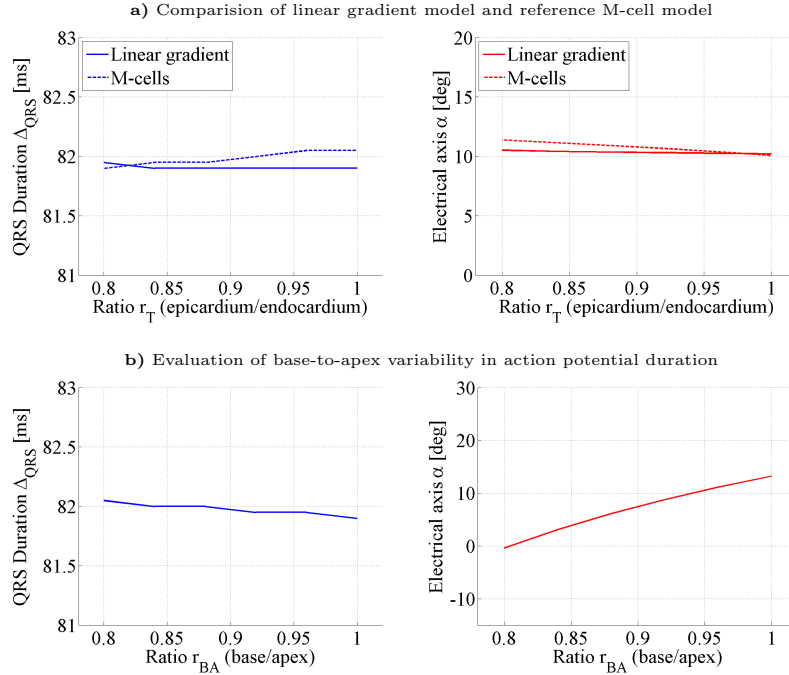


Figure 10: Influence of action potential duration (τ_{close}) on ECG features. **a)** The difference between the used linear gradient model and an M-cell model is negligible; the transmural ratio does not seem to significantly influence QRS duration and electrical axis. **b)** A downward gradient (base to apex) only causes slight variation in ECG features, showing that regional differences in τ_{close} cause low variation in ECG features. *See text for details.*

463 8.2 ms. The electrical axis varied by 30.8° as the electrical activation pattern
 464 was modified due to the anisotropic diffusivity. However, that variation was
 465 still within the clinical range.

466 Altogether, these finding – linear dependence of Δ_{QRS} on myocardial
 467 diffusion, and linear dependence of α on LV/RV endocardial diffusion – con-
 468 firmed the assumptions made in Sec. 2.2.3 and justified the selection of the
 469 two features for the estimation of cardiac electrical diffusivity. However, as
 470 expected, the experiments also showed that multiple combinations of c_{Myo} ,
 471 c_{LV} and c_{RV} can yield the same set of ECG features. The resulting uncer-
 472 tainty of diffusion parameters given a set of ECG features will be quantified
 473 in Sec. 3.2.2. Because variation for different APD distributions and fiber el-
 474 evation angles was found to be relatively low or cannot be directly measured

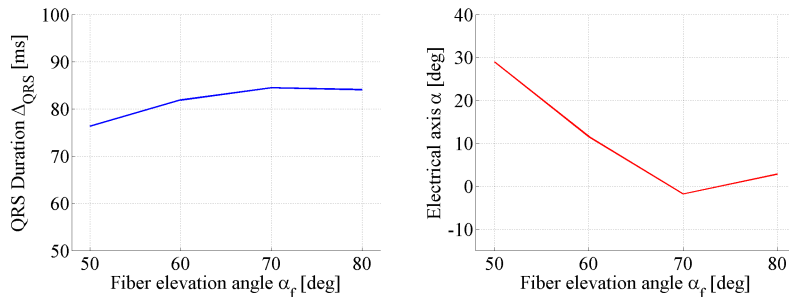


Figure 11: Influence of fiber elevation angle α_f within physiological range on ECG features (the anatomical models are generated with α_f on the endocardium and $-\alpha_f$ on the epicardium.)

475 in-vivo at the time of this study, we focused on the estimation of diffusion
 476 coefficients, keeping the other parameters at their nominal value.

477 3.1.3. Analysis of Computational Efficiency

478 Using one representative patient case, computational efficiency was ana-
 479 lyzed on a system with a 16-core Intel Xeon 64-bit CPU at 2.4 GHz and an
 480 NVIDIA GeForce GTX 580 graphics card. As described by Zettinig et al.
 481 (2013b), the computational times of image preparation and anatomical model
 482 creation, which has to be computed only once per patient, amounted to a
 483 total of 81.2 seconds. Table 4 reports the runtimes of the LBM-EP algo-
 484 rithm for a full heart cycle on differently spaced Cartesian grids (Georgescu
 485 et al., 2013). The projection of the extracellular potentials to the torso and
 486 the calculations of the ECG traces are simple matrix operations. Hence, the
 487 evaluation of the complete forward model could be done in less than 3 sec-
 488 onds for a grid with an isotropic resolution of 1.5 mm. As the evaluation of
 489 a polynomial function is almost immediate, the estimation of cardiac diffu-
 490 sivity required less than 10 seconds because of the three forward runs for the
 purpose of normalization.

Table 4: Full heart cycle runtimes of the LBM-EP algorithm for different grid spac-
 ings (Georgescu et al., 2013)

Grid Spacing	GPU Runtime
1.5 mm	2.8 sec.
0.7 mm	21.7 sec.

491

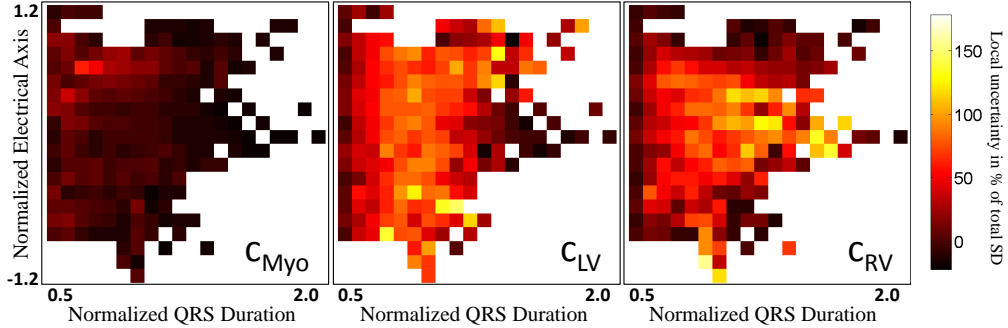


Figure 12: Estimated diffusion standard deviation (SD) in % of total SD for known electrical axis and QRS duration. The highest uncertainty is found in the healthy range of parameters (center of plots).

492 3.2. Evaluation of the Proposed Data-Driven Estimation Framework

493 3.2.1. Experimental Protocol

494 In this study, datasets of 19 patients with dilated cardiomyopathy (DCM)
 495 and a QRS duration of at least 120 ms were used. For all of them, an anatomical
 496 model was created based on cine magnetic resonance images (MRI) as
 497 described in Sec. 2.1. Thereafter, 500 EP simulations were computed for
 498 each patient on a 1.5 mm-isotropic Cartesian grid, accounting for a total of
 499 9,500 forward model runs. Diffusivity coefficients were uniformly sampled
 500 between $50 \text{ mm}^2/\text{s}$ and $5,000 \text{ mm}^2/\text{s}$ under the constraints $c_{Myo} \leq c_{LV}$ and
 501 $c_{Myo} \leq c_{RV}$.

502 3.2.2. Uncertainty Analysis in Cardiac Diffusion Parameters

503 Before training the regression model, the intrinsic uncertainty of the ECG
 504 inverse problem under our forward model was quantified using the entire syn-
 505 thetic EP database (9,500 simulations). To minimize the effects of geometry,
 506 the analysis was conducted with normalized ECG parameters. All computed
 507 $(\overline{\Delta_{QRS}}, \bar{\alpha})$ tuples were grouped in 20×20 bins, and for each bin, the local
 508 standard deviation of the diffusion coefficients c_{Myo} , c_{LV} and c_{RV} was calcu-
 509 lated. Table 5 reports the total standard deviation in the entire dataset, the
 510 average local standard deviation, and the uncertainty defined as their ratio.
 511 As illustrated in Fig. 12, which shows the uncertainty for each bin, the high-
 512 est variation can be found in the healthy range of QRS duration and electrical
 513 axis (up to 180%). The reported high uncertainties, especially for c_{LV} and
 514 c_{RV} , reflects the ill-posed nature of the ECG inverse problem if only QRS

515 duration and electrical axis are employed to personalize the model. That in-
 516 formation will be useful when evaluating the accuracy of the personalization
 517 techniques in the next sections.

Table 5: Total, and average local (bin-wise) standard deviation, and the uncertainty de-
 fined as their ratio for all three diffusion coefficients.

	c_{Myo}	c_{LV}	c_{RV}
Total SD (mm^2/s)	1,482	1,095	1,191
Avg. local SD (mm^2/s)	191	556	537
Uncertainty	12.9%	50.7%	45.1%

518 3.2.3. Evaluation on Synthetic Data

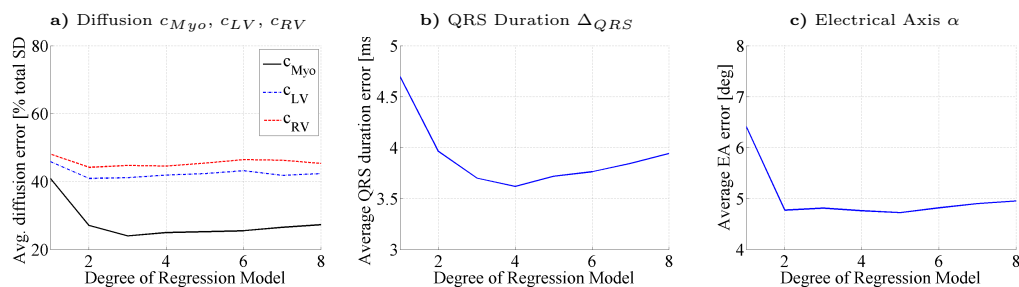


Figure 13: Analysis of polynomial regression degree on prediction accuracy. Average test-
 ing errors of leave-one-patient-out cross-validation in **a)** diffusion space, **b)** QRS duration
 Δ_{QRS} , and **c)** electrical axis α . The regression model overfits at degrees ≥ 4 , as the
 prediction errors in ECG feature space increase again.

519 The proposed machine-learning personalization procedure was evaluated
 520 using a leave-one-patient-out cross-validation on the database, i.e. the re-
 521 gression models were trained using a subset of 18 patients and tested with
 522 the remaining one, for each of the 19 DCM cases respectively. Next, the
 523 average testing errors in the diffusion (parameter) space were calculated. To
 524 evaluate the accuracy of the regression model in the observable space of ECG
 525 parameters, Δ_{QRS} and α were computed according to the estimated diffusiv-
 526 ity parameters and quantitatively compared with the known ground truth.
 527 In order to analyze the required dimensionality of the polynomial regression
 528 model, a cross-validation procedure with regression degrees ranging from 1
 529 to 8 was performed. While linear or quadratic regression models failed to

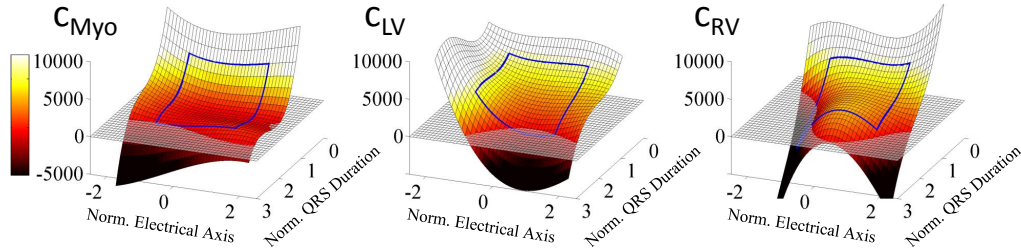


Figure 14: Regressed surface for each diffusion parameter. The highlighted area (blue contour) corresponds to the normalized parameter range in Fig. 12. As expected, all three parameters are dependent on the QRS duration, and the electrical axis has little effect on c_{Myo} . Right axis deviations correctly lead to $c_{RV} < c_{LV}$, and vice versa. Beyond the normalization range, all polynomials may drop below zero (non-physical diffusivity) for certain parameter combinations. *See text for details.*

530 capture the ECG problem, as shown in Fig. 13, the model started to overfit
 531 at degrees higher than four, leading to again increasing prediction errors in
 532 ECG space. Thus, the best option is to use cubic regression.

533 In Tab. 6, the final regression coefficients $\beta_{i,j}$ according to Eq. 13, trained
 534 using the entire synthetic dataset, are given. In Fig. 14, the regressed sur-
 535 faces for each of the diffusion parameters are visualized. Myocardial diffusion
 536 c_{Myo} is, as expected, almost exclusively dependent on the QRS duration. In-
 537 terestingly, for long QRS durations ($\Delta_{QRS} \geq 1.7$), c_{RV} decreases for left axis
 538 deviations while still remaining higher than c_{LV} . Nevertheless, axis devia-
 539 tions on both sides are correctly captured, i.e. c_{RV} is clearly smaller than
 540 c_{LV} for deviations to the right, and c_{LV} is slightly smaller than c_{RV} for de-
 541 viations to the left. Beyond the highlighted normalization range, all three
 542 polynomials may drop below zero for certain combinations of QRS duration
 543 and electrical axis, yielding non-physical diffusion coefficients. A test after
 544 the prediction can be used to discard these cases. The errors in estimated
 545 diffusion reported in Tab. 7 were obtained using this model. The relative
 546 errors in % of the total standard deviation of the dataset were in the same
 547 range as the estimated uncertainty of the inverse problem (Sec. 3.2.2). The
 548 proposed regression model was thus able to predict up to the intrinsic un-
 549 certainty of the problem. Furthermore, prediction errors were significantly
 550 higher when no normalization was applied, as illustrated in Tab. 7, suggest-
 551 ing the proposed model-based normalization procedure was able to partially
 552 compensate for inter-patient geometry variability.

553 In addition to the diffusion parameters used in the forward model (c_{Myo} ,

554 c_{LV} and c_{RV}), we tested how well the ratio between c_{LV}/c_{RV} can be recon-
555 structed. Low prediction errors as listed in Tab. 7 were expected in light of
556 the experiments carried out in the previous section, which showed a linear
557 dependency of the electrical axis on c_{LV} when $c_{LV} + c_{RV}$ is kept constant.
558 However, the ratio alone is not sufficient for a complete model personalization
559 as the two values are needed.

Table 6: Learned regression coefficients $\beta_{i,j}$ rounded to 5th decimal position (Eq. 13, Δ_{QRS} and α to be given in seconds and radians, respectively). Full double-precision coefficients are available from supplementary material.

i	j	g_{Myo}	g_{LV}	g_{RV}
3	0	-4397.72303	-2224.80345	-2372.28284
2	1	947.10794	-217.53110	-111.83210
2	0	20619.61231	8748.58856	10659.98012
1	2	-339.47629	884.12470	-1849.11181
1	1	-2138.03900	-556.40419	1168.22703
1	0	-31323.62564	-13516.54750	-15836.31353
0	3	-125.51341	-238.65933	513.65578
0	2	688.60647	-215.02537	1004.99532
0	1	1070.11229	585.32229	-900.63528
0	0	15662.21934	9621.33473	10454.35893

Table 7: Diffusion space prediction errors on the synthetic dataset, absolute in mm^2/s and relative in % of the total standard deviation. In addition to the three parameters, also the ratio between c_{LV} and c_{RV} was tested.

	c_{Myo}	c_{LV}	c_{RV}	c_{LV}/c_{RV}
With Normalization	356	451	533	
	24.0%	41.2%	44.7%	21.3%
Without normalization	571	540	597	
	38.5%	49.3%	50.0%	23.9%

560 *Comparison Against Nominal Values.* Table 8 reports the average absolute
561 errors in ECG feature space for forward model simulations with nominal dif-
562 fusion parameters from literature and parameters obtained with the proposed
563 regression framework. Likewise, the error distributions are shown in Fig. 15.
564 Calibrated simulations using our framework were not only in the range of

565 clinical variability but also significantly (t-test p-value < 0.001) more pre-
 566 cise than those obtained with nominal diffusivity values. In addition, our
 567 predictions were on average centered around the ground truth QRS dura-
 568 tion (average bias: +0.7 ms), the Δ_{QRS} calculated with default parameters
 569 was on average 28.9 ms too short. As the default parameters correspond to
 570 healthy physiology whereas conduction abnormalities cause prolonged QRS
 571 durations, this result was expected. Using our diffusion estimation frame-
 572 work may thus be preferable to using nominal parameters when only ECG
 573 is available.

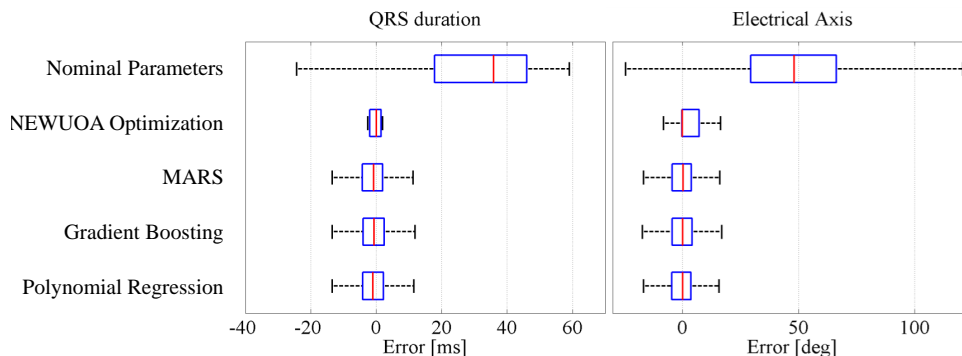


Figure 15: QRS duration and electrical axis error distributions for ECG simulations with nominal, NEWUOA-estimated and predicted diffusivity parameters using a machine learning technique. On each box, the central mark is the median, the edges of the box are the quartiles, and the whiskers extend to the most extreme data points not considered outliers. The range between the whiskers covers approximately 99.3% of the normally distributed data.

Table 8: Average absolute ECG feature space errors for ECG simulations with nominal, NEWUOA-estimated and predicted diffusivity parameters using a machine learning technique.

Diffusivity	Δ_{QRS} [ms]		α [deg]	
Nominal Parameters	33.7	± 15.7	53.2	± 33.8
NEWUOA Optimization	7.4	± 11.3	16.1	± 31.4
MARS	4.6	± 5.1	9.8	± 23.0
Gradient Boosting	4.9	± 5.8	9.5	± 19.8
Polynomial Regression	4.8	± 6.0	8.9	± 19.7

574 *Comparison against Alternative Machine Learning Techniques.* In this study,
575 the predictive power of the proposed polynomial regression framework is com-
576 pared against two non-parametric non-linear methods: multivariate adaptive
577 regression splines (MARS) and Gradient Boosting, as described in Hastie
578 et al. (2009). The former, MARS, is a non-parametric regression method
579 with explicative capabilities, which intuitively extends linear regression by
580 fitting splines to the predictors to capture data non-linearities and variable
581 interactions. For our evaluation, the ARESLab toolbox (Jekabsons, 2011)
582 was used. Gradient Boosting, on the other hand, is based on an ensemble
583 of weak prediction models, in our case 100 decision trees (LSBoost function
584 of MATLAB). Tab. 9 lists the diffusion space errors for all tested machine
585 learning algorithms. The errors in ECG feature space can be found in Tab. 8
586 and Fig. 15. Both approaches yielded very similar diffusion error distribu-
587 tions compared to the proposed polynomial regression framework. Also the
588 error distributions of Δ_{QRS} and α obtained by MARS and Gradient Boosting
589 were similar to those obtained by polynomial regression.

Table 9: Diffusion space prediction errors on the synthetic dataset, relative in % of the total standard deviation, for the tested machine learning algorithms.

	c_{Myo}	c_{LV}	c_{RV}
MARS	23.2%	40.2%	43.7%
Gradient Boosting	24.3%	46.1%	49.1%
Polynomial Regression	24.0%	41.2%	44.7%

590 *Comparison against an Alternative Inverse-Problem Method.* We compared
591 the performance of the regression framework with a personalization approach
592 that is based on NEWUOA (Powell, 2006, 2008), a gradient-free inverse prob-
593 lem method. An algorithm similar to the approach proposed by Neumann
594 et al. (2014) is followed. The diffusion coefficients c_{Myo}^0 , c_{LV}^0 , c_{RV}^0 are initial-
595 ized with parameters associated with healthy EP (Tab. 2, configuration F_1).
596 The initial step size is set to 500 mm²/s. Cardiac diffusivity is then estimated
597 using NEWUOA such that both the QRS durations and the electrical axis
598 match:

$$(c_{Myo}^*, c_{LV}^*, c_{RV}^*) = \arg \min_{c_{Myo}, c_{LV}, c_{RV}} C [f_{\Delta_{QRS}, \alpha}(c_{Myo}, c_{LV}, c_{RV})] \quad (16)$$

599 Hereby, $f_{\Delta_{QRS}, \alpha}(\cdot)$ denotes the ECG features obtained by running the EP
600 forward model. In the cost function C , the values Δ_{QRS}^m and α^m are the

601 measured QRS duration and electrical axis, respectively, and the parameter
 602 $\lambda = 0.1$ accounts for the different orders of magnitude between QRS duration
 603 (in seconds) and electrical axis (in radians):

$$C(\Delta_{QRS}^i, \alpha^i) = |\Delta_{QRS}^m - \Delta_{QRS}^i| + \lambda |\alpha^m - \alpha^i| \quad (17)$$

604 As shown in Tab. 8, the errors in Δ_{QRS} and α calculated using the NEWUOA-
 605 personalized forward model were higher compared to the data-driven estima-
 606 tion framework. In addition, the obtained values for the electrical axis were
 607 less centered around the ground truth (average bias: 8.1°). Note that the
 608 table lists higher standard deviations (11.3 ms and 31.4° for Δ_{QRS} and α ,
 609 respectively) than Fig. 15 suggests because of numerous outliers (16%) in the
 610 NEWUOA predictions.

611 Because the optimizer was sensitive to local minima, NEWUOA typically
 612 required up to 50 iterations to converge to a stable optimum. Thus, the
 613 total time for optimization using a fully optimized version of the code would
 614 take up to 2.5 min, while our approach requires only 10 s to calculate the
 615 three forward simulations for the normalization. Our approach was therefore
 616 not only at least $15\times$ more computationally efficient but also yielded more
 617 predictive diffusivity parameters.

618 3.2.4. Evaluation on Real DCM Cases

619 Finally, we evaluated the machine-learning personalization with the clin-
 620 ical ECG data which were available for all 19 DCM cases. The trained re-
 621 gression models from the cross-validation were employed to estimate diffusion
 622 coefficients based on measured QRS duration and electrical axis. In three
 623 cases, the prediction was not successful and yielded negative diffusivity for at
 624 least one of the diffusion parameters because the measured electrical axis was
 625 outside the normalization range. These cases are easily identifiable and could
 626 therefore be processed using other approaches if needed. For the remaining
 627 16 patients, plausible diffusion coefficients (between 141 and $582 \text{ mm}^2/\text{s}$ for
 628 c_{Myo} , and between 678 and $2769 \text{ mm}^2/\text{s}$ for c_{LV} and c_{RV}) were estimated.
 629 Table 10 reports the average absolute errors between clinical ground truth
 630 and ECG features obtained with forward model computations using the es-
 631 timated diffusion parameters for the remaining 16 patients. Figure 16 shows
 632 the obtained error distributions, indicating that the simulated QRS dura-
 633 tion was on average 18 ms too long, while the electrical axis was closely
 634 centered (average bias: 3.1°) around the measurements, both values being

635 within clinical acceptability as defined prior to the study. Finally, Fig. 17
 636 illustrates the simulated ECG chest leads overlaid on the measured ones for
 637 one representative patient.

Table 10: Average absolute ECG feature space errors for ECG computations with regression-predicted diffusivity parameters from clinically measured ECG features.

Diffusivity	Δ_{QRS} [ms]	α [deg]
Regression-based Prediction	18.7 \pm 16.2	6.5 \pm 7.6

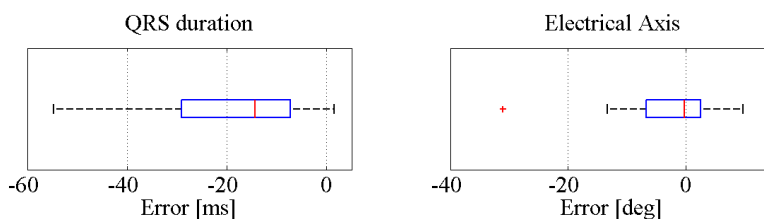


Figure 16: QRS duration and electrical axis error distributions for ECG simulations diffusivity parameters estimated from clinical ground truth measurements.

638 4. Discussion and Conclusions

639 4.1. Discussion

640 In this manuscript, we described a data-driven method for the person-
 641 alization of a cardiac electrophysiology model from ECG features. As sup-
 642 ported by reported results, the method achieves the same accuracy as tra-
 643 ditional inverse problem algorithm with the advantage of 1) being computa-
 644 tionally efficient (evaluation of a polynomial function is almost immediate)
 645 and 2) providing an estimate of parameter uncertainty, an additional variable
 646 that could be employed clinically.

647 While the anatomical model was obtained from cine MR images in this
 648 work, the approach is easily applicable to other modalities such as computed
 649 tomography (CT) or echocardiography (Zheng et al., 2008), provided the full
 650 bi-ventricular myocardium is visible. In this study, a synthetic model of fiber
 651 architecture was employed. Nevertheless, our framework is ready to incor-
 652 porate diffusion tensor imaging (DTI) data and thus remove this additional
 653 uncertainty. It will be important to quantify the difference between ex-vivo
 654 measurements (Helm et al., 2005) and in-vivo measurements in order to learn

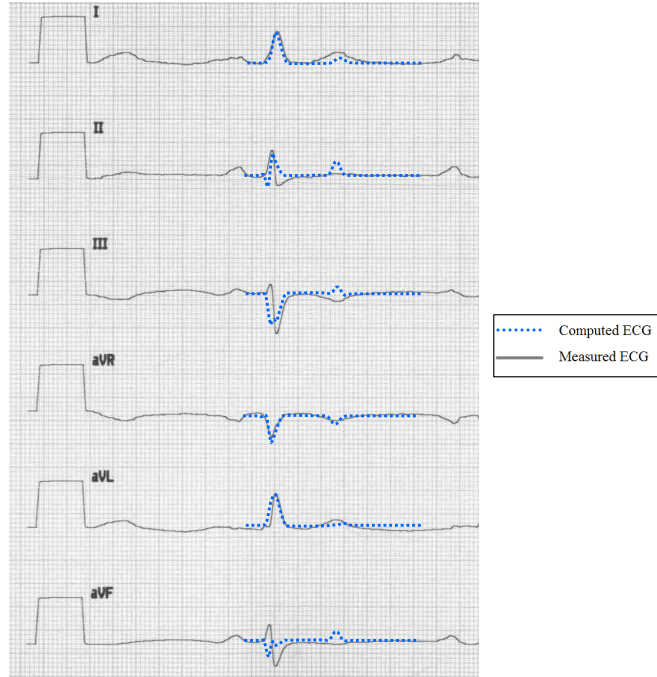


Figure 17: Clinically measured, and computed ECG chest leads after model estimation of cardiac diffusivity for one representative patient, showing promising agreement during cardiac depolarization, which we focused on in this study. For this case, obtained estimation errors amounted to 1.6 ms for the QRS duration and 0.5° for the electrical axis.

655 a data-driven model for patient-specific diffusivity predictions. Furthermore,
 656 it would be interesting to investigate how the variances of the atlas gener-
 657 ated by Lombaert et al. (2012) can be used for a sensitivity analysis and
 658 uncertainty estimation in our model. In any case, the tremendous progress
 659 achieved in in-vivo DTI will soon enable to use patient-specific data (Tous-
 660 saint et al., 2013). For the mapping of potentials onto the body surface, an
 661 atlas of torso geometry was employed as 3D images of patient upper body
 662 were not available. The manual registration of the atlas against 2D contours
 663 outlined in the three sagittal, axial and longitudinal planes was performed by
 664 an expert. It should be noted however that slight mis-registration would not
 665 have impacted the performance of the algorithm as ECG leads are known to
 666 be tolerant with respect to electrode placement (Sheppard et al., 2011). In
 667 addition, the overall methodology would work on more detailed atlases based
 668 on populations or on the patient-specific geometry directly.

669 This work was performed using a mono-domain EP model with the ac-
670 tion potential model proposed by Mitchell and Schaeffer (2003). It has been
671 shown by Boulakia et al. (2010) and Plank et al. (2013) that anisotropic
672 mono-domain models are able to preserve the essential ECG features, which
673 were used for the subsequent personalization, when compared to orthotropic
674 bi-domain models. It should be noted that the LBM-EP method can use any
675 mono-domain model like for instance the TenTusscher model. Furthermore,
676 the data-driven personalization algorithm is generic by design and can be
677 applied to any cell model, or any bi-domain or graph-based/Eikonal model
678 of cardiac electrophysiology, as far as the database can be computed in a
679 realistic amount of time. Moreover, our focus on cardiac depolarization al-
680 lowed decoupling the estimation of electrical diffusivity from repolarization
681 EP parameters and assuming a static heart. Unlike during the ST-T period,
682 the deformation of the myocardium due to cardiac motion has been shown
683 to be marginal during the QRS complex (Jiang et al., 2009). Also, the influ-
684 ence of the action potential duration on the ECG features used to estimate
685 electrical diffusion (QRS duration and electrical axis) was confirmed to be
686 negligible (Fig. 10).

687 In this study, the mapping of extracellular potentials from the heart to the
688 body surface relies on a boundary element approach. For our simulations, we
689 applied constant homogeneous isotropic conductivity in the torso, including
690 the chest cavity, thoracic cage, muscle tissue and skin. Minor sensitivity
691 on body surface potentials for different organ conductivities as observed in
692 the computational study by Geneser et al. (2008) justifies this assumption
693 for our purposes. We verified our BEM implementation with analytically
694 defined test cases and showed convergence with increasing mesh resolution
695 and order of Gaussian quadrature for the evaluation of integrals without
696 available closed-form formula (Fig. 8), suggesting that the uncertainty in
697 diffusivity parameters is not related to BEM numerical approximations but
698 rather intrinsic to the inverse ECG problem.

699 From the ECG traces, two features were derived, namely the QRS du-
700 ration Δ_{QRS} and the electrical axis α . Provided the choice of appropriate
701 electrode positions, experiments have shown that the electrical axis can be
702 computed on either the epicardium mesh or the torso mesh. A thorough eval-
703 uation of the influence of varying conductivities in the torso due to different
704 organs (Geneser et al., 2008) onto the electrical axis measured at the torso
705 would be interesting for future works. Regarding the computation of the QRS
706 duration, we observed in our experiments a close match between the dura-

707 tion computed directly using the LBM-EP results as described in Sec. 2.2.3
708 and the width of the QRS complex in the computed ECG leads. This work
709 is based on the assumption that these two features are sufficient to explain
710 various EP patterns. In addition, the selected features are commonly avail-
711 able from clinical ECG traces and clinical reports, and would therefore allow
712 the estimation of diffusion coefficients with little effort in clinical routine.
713 The proposed method could therefore constitute a first model personaliza-
714 tion step when no dense EP data is available, and would also provide more
715 accurate results compared to generic parameters, as suggested by our exper-
716 iments. Yet, the proposed framework is generic and allows the integration
717 of an arbitrary number of features. If invasive or BSM measurements are
718 available, more regional features could be used to facilitate the estimation of
719 model parameters.

720 The use of QRS duration and electrical axis was further supported by
721 our parameter analysis. It has been shown in this paper that QRS dura-
722 tion varies linearly with myocardial diffusion (c_{Myo}), while the electrical axis
723 varies linearly with increasing left endocardial diffusion (c_{LV}) when the right
724 endocardial diffusion (c_{RV}) is decreased at the same time such that their sum
725 is constant. However, these relationships are not decoupled as each diffusion
726 parameter has influence on both features, which contributes to the quanti-
727 fied uncertainty of the inverse problem: different diffusivity configurations
728 can lead to the same ECG parameters. In particular, we showed that left
729 and right endocardial diffusivity are subject to broad variations, especially
730 in the region of healthy EP (Fig. 12). Clearly demonstrating the ill-posed
731 nature of the inverse ECG problem under the assumptions of our EP model,
732 the reported uncertainties constitute, to the best of our knowledge, the first
733 estimates of the optimal bound in accuracy for any inverse problem to esti-
734 mate myocardium diffusion that rely on Δ_{QRS} and α only. We expect the
735 uncertainty to decrease as more clinical features are considered. This study
736 is thus subject to future work.

737 The main contribution of this work is our novel data-driven framework
738 to estimate cardiac diffusion parameters. Instead of solving the inverse ECG
739 problem numerically, we proposed to employ statistical learning, and in par-
740 ticular multivariate polynomial regression, to learn the relationship between
741 ECG features and diffusivity. Compared to other statistical approaches, poly-
742 nomial regression has the advantage that the regression coefficients can be
743 given and the estimation of diffusion parameters is possible using a closed-
744 form formula. The personalization formula can therefore be shared between

745 research groups. Error distributions obtained using multivariate adaptive re-
746 gression splines (MARS) and Gradient Boosting were, as reported in Tab. 8,
747 similar to those obtained by polynomial regression. An evaluation of the
748 required polynomial degree revealed that the model starts over-fitting at de-
749 gree 4 (Fig. 13). We therefore use cubic multivariate regression and report
750 the final coefficients in Tab. 6.

751 A key aspect of the approach is the model-based normalization of EP fea-
752 tures to indirectly incorporate geometric information in the statistical model.
753 The strategy consists in scouting the space of Δ_{QRS} and α for a given patient
754 by running three forward simulations with diffusion parameters relating to
755 healthy EP, and left and right bundle branch block scenarios. As a result,
756 although not directly based on anatomical or physiological features such as
757 heart size or strength of myocardial contraction, we were able to show that
758 the normalization scheme compensates for patient geometry and significantly
759 improves prediction results (Tab. 7). For an unseen patient, three forward
760 model runs are needed (computed in about 10 seconds using LBM-EP), which
761 is still acceptable in a clinical setting but also far less than in conventional
762 inverse-problem algorithms, which require often numerous model evaluations
763 to converge.

764 As expected, the regressed surfaces of the three diffusivities trained from
765 a database of 9,500 simulations (Fig. 14) showed dependance of the QRS du-
766 ration on all three diffusion parameters. For right axis deviations, the right
767 endocardial diffusion was consistently lower than the left one, and vice versa.
768 However, all three polynomials may yield negative, non-physical diffusion co-
769 efficients for certain parameter combinations. As not reported, preliminary
770 experiments have shown, regressing the logarithm of the diffusion coefficients
771 ($\log c$ instead of c directly) could help in overcoming this limitation. Future
772 work will therefore consider the investigation of regression methods with
773 better generalization on unseen data. Prediction errors in diffusion space
774 (leave-one-patient-out) were in the range of the estimated intrinsic uncer-
775 tainty of the problem, especially for left and right endocardial diffusivity
776 (Tab. 7). Only for myocardial diffusivity, the prediction was slightly worse
777 (24% of total std. dev.) compared to the uncertainty (13% of total std. dev.)
778 One reason for this result could be the dependence of myocardial diffusivity
779 on both ECG features for diffusions of less than 400 mm²/s (Fig 9a).

780 Yet, prediction errors in the ECG feature space for the synthetic dataset
781 obtained by running forward simulations using the calibrated EP model were
782 significantly better than those obtained by using nominal diffusion param-

783 eters from literature (Tab. 8). Furthermore, a comparison with an estima-
784 tion algorithm based on the gradient-free inverse problem method NEWUOA
785 showed that our method performs better (prediction errors in QRS duration
786 and electrical axis were 54% and 80% higher, respectively) while being im-
787 mediate to compute and providing uncertainty estimates. Altogether, the
788 application of the proposed data-driven framework may thus be preferential
789 to traditional approaches when only ECG data are available.

790 Finally, an evaluation with clinically measured ECG features was con-
791 ducted on all 19 patient cases. The model was successfully fitted in 16 cases
792 out of 19 (84%), with promising prediction errors of 18.7 ± 16.2 ms for Δ_{QRS}
793 and $6.5 \pm 7.6^\circ$ for α , within clinical acceptability. The model could not be
794 personalized in three cases as the measured electrical axis was outside the
795 normalization range. Similar results were obtained in not reported, prelim-
796 inary experiments on regressing the logarithm of the diffusion coefficients,
797 as mentioned above. However, the cases that failed could not be recovered
798 either, suggesting a more intrinsic difference between patient physiology and
799 our model, for instance a potential line of blocks, etc. A more realistic in-
800 corporation of geometrical features might improve the success rate and avoid
801 such inconsistencies. Further investigation is needed for an optimal regression
802 model.

803 The results previously published by Zettinig et al. (2013a) showed smaller
804 overall errors in Δ_{QRS} and α , potentially because 1) fewer patients with a
805 smaller range of anatomical and physiological variation were used, and 2) the
806 original framework (multivariate polynomial regression of degree 7) might
807 have been overfitting as shown in Fig. 13. In addition, estimation errors
808 without normalization were better than previously reported, possibly due to
809 the updated torso registration technique.

810 *4.2. Perspectives*

811 In this work, only the cardiac anatomy model was generated based on pa-
812 tient data. Despite the contour-based registration, torso geometry was based
813 on an atlas and does not entirely reflect patient-specific anatomy. Also the
814 boundary element mapping of potentials assumes constant conductivity, ne-
815 glecting thoracic organs and different tissue types. Future work could thus
816 improve the anatomical model by incorporating more imaging data from the
817 heart to the body surface and model the different tissues in the torso in-
818 dependently (lung, bones, muscles). Furthermore, instead of the proposed
819 normalization technique, explicitly integrating geometrical features directly

820 into the regression framework could potentially better cope with anatomical
821 variability. In addition, the framework could be extended by using an elec-
822 tromechanical model of the heart (Zettinig et al., 2013b) to cope with the
823 influence of cardiac motion on the ECG. A comprehensive study is needed
824 though to quantify that aspect and properly consider it into the estimation
825 process. There are indeed no studies available to clarify how much mo-
826 tion happens during the fast depolarization of the heart, to the best of our
827 knowledge. Even though, a dynamic model would have great benefit when
828 estimated cardiac repolarization features like action potential duration.

829 As the uncertainty in diffusion parameters given QRS duration and elec-
830 trical axis is high, the integration of more ECG features could improve es-
831 timation precision and increase the success rate of the approach. Similarly,
832 more ECG features may potentially allow the estimation of more local diffu-
833 sion coefficients, rendering the estimation of regional diffusivity distributions
834 possible.

835 Finally, refining the forward model, in particular regarding cardiac elec-
836 trophysiology, might lead to future extensions of our framework. More com-
837 plex biophysical bi-domain models, integration of atrial geometry, more re-
838 fined activation patterns, and coupling with mechanical models could po-
839 tentially increase the predictive power of the framework and are subject to
840 future work.

841 *4.3. Conclusion*

842 We have shown in this paper that the estimation of patient-specific car-
843 diac diffusion parameters from standard 12-lead ECG measurements using
844 machine learning techniques is possible, up to the intrinsic uncertainty of
845 the problem. Based on QRS duration and electrical axis as ECG features,
846 a data-driven regression model was trained and used to predict diffusivity
847 parameters for left and right endocardium (mimicking the fast conducting
848 Purkinje system), and the bulk myocardium tissue. Under the assumptions
849 of our forward model, the prediction errors were in the range of the under-
850 lying uncertainty in diffusivity, which we empirically quantified for the first
851 time to the best of our knowledge. We evaluated the framework both on the
852 synthetic dataset and on clinical measurements using a leave-one-patient-
853 out cross-validation and computed the error in ECG feature space using
854 forward simulations with estimated diffusion parameters. Significant im-
855 provement with respect to nominal diffusivity values, which relate to healthy
856 electrophysiology, were obtained. We also conducted a comparison with a

857 NEWUOA-based personalization approach, finding overall superior predic-
858 tive power. Therefore, our framework can provide good preliminary person-
859 alization, prior to more refined estimation if invasive or BSM measurements
860 are available.

861 **References**

862 Aliev, R.R., Panfilov, A.V., 1996. A simple two-variable model of cardiac
863 excitation. *Chaos, Solitons & Fractals* 7, 293–301.

864 Arsigny, V., Fillard, P., Pennec, X., Ayache, N., 2006. Log-euclidean metrics
865 for fast and simple calculus on diffusion tensors. *Magnetic resonance in*
866 *medicine* 56, 411–421.

867 Auricchio, A., Prinzen, F.W., et al., 2011. Non-responders to cardiac resyn-
868 chronization therapy: the magnitude of the problem and the issues. *Cir-*
869 *culation Journal: Official Journal of the Japanese Circulation Society* 75,
870 521–527.

871 Barr, R.C., Ramsey, M., Spach, M.S., 1977. Relating epicardial to body
872 surface potential distributions by means of transfer coefficients based on
873 geometry measurements. *Biomedical Engineering, IEEE Transactions on*
874 *24*, 1–11.

875 Bayer, J., Blake, R., Plank, G., Trayanova, N., 2012. A novel rule-based
876 algorithm for assigning myocardial fiber orientation to computational heart
877 models. *Annals of biomedical engineering* 40, 2243–2254.

878 Boulakia, M., Cazeau, S., Fernández, M.A., Gerbeau, J.F., Zenzemi, N.,
879 2010. Mathematical modeling of electrocardiograms: a numerical study.
880 *Annals of biomedical engineering* 38, 1071–1097.

881 Bourgault, Y., Coudière, Y., Pierre, C., 2009. Existence and uniqueness
882 of the solution for the bidomain model used in cardiac electrophysiology.
883 *Nonlinear analysis: Real world applications* 10, 458–482.

884 Chhay, M., Coudière, Y., Turpault, R., 2012. How to compute the extracel-
885 lular potential in electrocardiology from an extended monodomain model.
886 *Research Report RR-7916. INRIA.*

- 887 Chinchapatnam, P., Rhode, K.S., Ginks, M., Rinaldi, C.A., Lambiase, P.,
888 Razavi, R., Arridge, S., Sermesant, M., 2008. Model-based imaging of
889 cardiac apparent conductivity and local conduction velocity for diagnosis
890 and planning of therapy. *Medical Imaging, IEEE Transactions on* 27,
891 1631–1642.
- 892 Chung, E.K., 1989. *Comprehensive electrocardiography: Theory and prac-*
893 *tice in health and disease.* *JAMA: The Journal of the American Medical*
894 *Association* 262, 571–571.
- 895 Clayton, R., Bernus, O., Cherry, E., Dierckx, H., Fenton, F., Mirabella, L.,
896 Panfilov, A., Sachse, F., Seemann, G., Zhang, H., 2011. Models of cardiac
897 tissue electrophysiology: progress, challenges and open questions. *Progress*
898 *in biophysics and molecular biology* 104, 22–48.
- 899 Clayton, R., Panfilov, A., 2008. A guide to modelling cardiac electrical
900 activity in anatomically detailed ventricles. *Progress in biophysics and*
901 *molecular biology* 96, 19–43.
- 902 Coudière, Y., Pierre, C., 2006. Stability and convergence of a finite vol-
903 ume method for two systems of reaction-diffusion equations in electro-
904 cardiology. *Nonlinear analysis: Real world applications* 7, 916–935.
- 905 Dössel, O., Krueger, M.W., Weber, F.M., Schilling, C., Schulze, W.H., See-
906 mann, G., 2011. A framework for personalization of computational mod-
907 els of the human atria, in: *Engineering in Medicine and Biology Society,*
908 *EMBC, 2011 Annual International Conference of the IEEE, IEEE.* pp.
909 4324–4328.
- 910 Fenton, F., Karma, A., 1998. Vortex dynamics in three-dimensional contin-
911 uous myocardium with fiber rotation: filament instability and fibrillation.
912 *Chaos: An Interdisciplinary Journal of Nonlinear Science* 8, 20–47.
- 913 Ferri, F., 2013. *Ferri’s Clinical Advisor 2013.* Elsevier Mosby.
- 914 FitzHugh, R., 1961. Impulses and physiological states in theoretical models
915 of nerve membrane. *Biophysical journal* 1, 445–466.
- 916 Franzone, P.C., Guerri, L., Rovida, S., 1990. Wavefront propagation in an
917 activation model of the anisotropic cardiac tissue: asymptotic analysis and
918 numerical simulations. *Journal of mathematical biology* 28, 121–176.

- 919 Geneser, S.E., Kirby, R.M., MacLeod, R.S., 2008. Application of stochastic
920 finite element methods to study the sensitivity of ecg forward modeling
921 to organ conductivity. *Biomedical Engineering, IEEE Transactions on* 55,
922 31–40.
- 923 Georgescu, B., Rapaka, S., Mansi, T., Zettinig, O., Kamen, A., Comaniciu,
924 D., 2013. Towards real-time cardiac electrophysiology computations using
925 GP-GPU Lattice-Boltzmann method, in: *Proceedings of 6th International
926 Workshop on High Performance Computing for Biomedical Image Analysis
927 (HPC-MICCAI)*. Springer.
- 928 Glukhov, A.V., Fedorov, V.V., Lou, Q., Ravikumar, V.K., Kalish, P.W.,
929 Schuessler, R.B., Moazami, N., Efimov, I.R., 2010. Transmural dispersion
930 of repolarization in failing and nonfailing human ventricle. *Circulation
931 research* 106, 981–991.
- 932 Guennebaud, G., Jacob, B., et al., 2010. Eigen v3.
933 <http://eigen.tuxfamily.org>.
- 934 Gulrajani, R.M., 1998. The forward and inverse problems of electrocardiog-
935 raphy. *Engineering in Medicine and Biology Magazine, IEEE* 17, 84–101.
- 936 Han, C., Pogwizd, S.M., Killingsworth, C.R., Zhou, Z., He, B., 2013. Nonin-
937 vasive cardiac activation imaging of ventricular arrhythmias during drug-
938 induced QT prolongation in the rabbit heart. *Heart Rhythm* 10, 1509–
939 1515.
- 940 Hastie, T., Tibshirani, R., Friedman, J., 2009. *The elements of statistical
941 learning: Data mining, inference, and prediction*. Springer.
- 942 Helm, P.A., Tseng, H.J., Younes, L., McVeigh, E.R., Winslow, R.L., 2005.
943 Ex vivo 3D diffusion tensor imaging and quantification of cardiac laminar
944 structure. *Magnetic resonance in medicine* 54, 850–859.
- 945 Hodgkin, A.L., Huxley, A.F., 1952. A quantitative description of membrane
946 current and its application to conduction and excitation in nerve. *The
947 Journal of physiology* 117, 500–544.
- 948 Jekabsons, G., 2011. ARESLab: Adaptive Regression Splines toolbox for
949 Matlab/Octave. URL: <http://www.cs.rtu.lv/jekabsons/>.

- 950 Jiang, M., Lv, J., Wang, C., Huang, W., Xia, L., Shou, G., 2011. A hybrid
951 model of maximum margin clustering method and support vector regres-
952 sion for solving the inverse ECG problem, in: *Computing in Cardiology*,
953 2011, IEEE. pp. 457–460.
- 954 Jiang, M., Xia, L., Shou, G., Wei, Q., Liu, F., Crozier, S., 2009. Effect
955 of cardiac motion on solution of the electrocardiography inverse problem.
956 *Biomedical Engineering, IEEE Transactions on* 56, 923–931.
- 957 Keener, J.P., Sneyd, J., 1998. *Mathematical physiology*. volume 8. Springer.
- 958 Konukoglu, E., Relan, J., Cilingir, U., Menze, B.H., Chinchapatnam, P., Ja-
959 didi, A., Cochet, H., Hocini, M., Delingette, H., Jaïs, P., et al., 2011. Ef-
960 ficient probabilistic model personalization integrating uncertainty on data
961 and parameters: Application to eikonal-diffusion models in cardiac elec-
962 trophysiology. *Progress in Biophysics and Molecular Biology* 107, 134–146.
- 963 Li, Z., Zhu, S., He, B., 2007. Solving the ECG forward problem by means
964 of a meshless finite element method. *Physics in medicine and biology* 52,
965 N287.
- 966 Liu, C., Eggen, M.D., Swingen, C.M., Iaizzo, P.A., He, B., 2012. Noninvasive
967 mapping of transmural potentials during activation in swine hearts from
968 body surface electrocardiograms. *Medical Imaging, IEEE Transactions on*
969 31, 1777–1785.
- 970 Lombaert, H., Peyrat, J., Croisille, P., Rapacchi, S., Fanton, L., Cheriet, F.,
971 Clarysse, P., Magnin, I., Delingette, H., Ayache, N., 2012. Human atlas
972 of the cardiac fiber architecture: Study on a healthy population. *Medical*
973 *Imaging, IEEE Transactions on* 31, 1436–1447.
- 974 Luo, C.h., Rudy, Y., 1991. A model of the ventricular cardiac action po-
975 tential. Depolarization, repolarization, and their interaction. *Circulation*
976 *research* 68, 1501–1526.
- 977 Marcus, G.M., Keung, E., Scheinman, M.M., 2013. The year in review of
978 cardiac electrophysiology. *Journal of the American College of Cardiology*
979 61, 772–782.
- 980 McMurray, J.J., Adamopoulos, S., Anker, S.D., Auricchio, A., Böhm, M.,
981 Dickstein, K., Falk, V., Filippatos, G., Fonseca, C., Gomez-Sanchez, M.A.,

- 982 et al., 2012. ESC guidelines for the diagnosis and treatment of acute and
983 chronic heart failure 2012. *European heart journal* 33, 1787–1847.
- 984 Mendis, S., Puska, P., Norrving, B., et al., 2011. Global atlas on cardiovas-
985 cular disease prevention and control. World Health Organization.
- 986 Michaels, L., Cadoret, R., 1967. Day-to-day variability in the normal elec-
987 trocardiogram. *British heart journal* 29, 913–919.
- 988 Mitchell, C.C., Schaeffer, D.G., 2003. A two-current model for the dynamics
989 of cardiac membrane. *Bulletin of mathematical biology* 65, 767–793.
- 990 Modre, R., Tilg, B., Fischer, G., Wach, P., 2002. Noninvasive myocardial
991 activation time imaging: a novel inverse algorithm applied to clinical eeg
992 mapping data. *Biomedical Engineering, IEEE Transactions on* 49, 1153–
993 1161.
- 994 Moireau, P., 2008. Assimilation de données par filtrage pour les systèmes hy-
995 perboliques du second ordre-Applications à la mécanique cardiaque. Ph.D.
996 thesis. Ecole Doctorale de l’Ecole Polytechnique, Paris.
- 997 Neumann, D., Mansi, T., Grbic, S., Voigt, I., Georgescu, B., Kayvanpour,
998 E., Amr, A., Sedaghat-Hamedani, F., Haas, J., Katus, H., Meder, B.,
999 Hornegger, J., Kamen, A., Comaniciu, D., 2014. Automatic image-to-
1000 model framework for patient-specific electromechanical modeling of the
1001 heart, in: *IEEE International Symposium on Biomedical Imaging 2014* (in
1002 press).
- 1003 Noble, D., 1962. A modification of the Hodgkin-Huxley equations applica-
1004 ble to Purkinje fibre action and pacemaker potentials. *The Journal of*
1005 *Physiology* 160, 317–352.
- 1006 Noble, D., Varghese, A., Kohl, P., Noble, P., 1998. Improved guinea-pig ven-
1007 tricular cell model incorporating a diadic space, IKr and IKs, and length-
1008 and tension-dependent processes. *The Canadian journal of cardiology* 14,
1009 123.
- 1010 Pernod, E., Sermesant, M., Konukoglu, E., Relan, J., Delingette, H., Ayache,
1011 N., 2011. A multi-front eikonal model of cardiac electrophysiology for
1012 interactive simulation of radio-frequency ablation. *Computers & Graphics*
1013 35, 431–440.

- 1014 Plank, G., Costa, C.M., Prassl, A.J., 2013. Parametrization strategies for
1015 matching activation sequences in models of ventricular electrophysiology,
1016 in: Engineering in Medicine and Biology Society (EMBC), 2013 35th An-
1017 nual International Conference of the IEEE, IEEE. pp. 1534–1537.
- 1018 Potse, M., Dubé, B., Richer, J., Vinet, A., Gulrajani, R.M., 2006. A com-
1019 parison of monodomain and bidomain reaction-diffusion models for action
1020 potential propagation in the human heart. *Biomedical Engineering, IEEE*
1021 *Transactions on* 53, 2425–2435.
- 1022 Potse, M., Dubé, B., Vinet, A., 2009. Cardiac anisotropy in boundary-
1023 element models for the electrocardiogram. *Medical & biological engineering*
1024 *& computing* 47, 719–729.
- 1025 Powell, M., 2006. The NEWUOA software for unconstrained optimization
1026 without derivatives, in: Pillo, G., Roma, M. (Eds.), *Large-Scale Nonlinear*
1027 *Optimization*. Springer US. volume 83 of *Nonconvex Optimization and Its*
1028 *Applications*, pp. 255–297.
- 1029 Powell, M.J., 2008. Developments of NEWUOA for minimization without
1030 derivatives. *IMA Journal of Numerical Analysis* 28, 649–664.
- 1031 Prakosa, A., Sermesant, M., Allain, P., Villain, N., Rinaldi, C., Rhode, K.,
1032 Razavi, R., Delingette, H., Ayache, N., 2013. Cardiac electrophysiological
1033 activation pattern estimation from images using a patient-specific database
1034 of synthetic image sequences. *Biomedical Engineering, IEEE Transactions*
1035 *on* 61, 235–245.
- 1036 Rapaka, S., Mansi, T., Georgescu, B., Pop, M., Wright, G.A., Kamen, A.,
1037 Comaniciu, D., 2012. LBM-EP: Lattice-Boltzmann method for fast cardiac
1038 electrophysiology simulation from 3D images, in: *Medical Image Comput-*
1039 *ing and Computer-Assisted Intervention—MICCAI 2012*. Springer LNCS.
1040 volume 7511, pp. 33–40.
- 1041 Relan, J., Chinchapatnam, P., Sermesant, M., Rhode, K., Ginks, M.,
1042 Delingette, H., Rinaldi, C.A., Razavi, R., Ayache, N., 2011. Coupled per-
1043 sonalization of cardiac electrophysiology models for prediction of ischaemic
1044 ventricular tachycardia. *Interface Focus* 1, 396–407.
- 1045 Roger, V.L., Go, A.S., Lloyd-Jones, D.M., Benjamin, E.J., Berry, J.D., Bor-
1046 den, W.B., Bravata, D.M., Dai, S., Ford, E.S., Fox, C.S., et al., 2012.

- 1047 Executive summary: heart disease and stroke statistics–2012 update: a
1048 report from the american heart association. *Circulation* 125, 188–197.
- 1049 Sermesant, M., Billet, F., Chabiniok, R., Mansi, T., Chinchapatnam, P.,
1050 Moireau, P., Peyrat, J.M., Rhode, K., Ginks, M., Lambiase, P., et al.,
1051 2009. Personalised electromechanical model of the heart for the prediction
1052 of the acute effects of cardiac resynchronisation therapy, in: *Functional
1053 Imaging and Modeling of the Heart*. Springer LNCS. volume 5528, pp.
1054 239–248.
- 1055 Sermesant, M., Konukoglu, E., Delingette, H., Coudière, Y., Chinchapatnam,
1056 P., Rhode, K.S., Razavi, R., Ayache, N., 2007. An anisotropic multi-front
1057 fast marching method for real-time simulation of cardiac electrophysiology,
1058 in: *Functional Imaging and Modeling of the Heart*. Springer LNCS. volume
1059 4466, pp. 160–169.
- 1060 Sethian, J.A., 1999. *Level set methods and fast marching methods: evolving
1061 interfaces in computational geometry, fluid mechanics, computer vision,
1062 and materials science*. volume 3. Cambridge university press.
- 1063 Sheppard, J.P., Barker, T.A., Ranasinghe, A.M., Clutton-Brock, T.H., Fren-
1064 neaux, M.P., Parkes, M.J., 2011. Does modifying electrode placement of
1065 the 12 lead ecg matter in healthy subjects? *International journal of car-
1066 diology* 152, 184–191.
- 1067 Shou, G., Xia, L., Jiang, M., Wei, Q., Liu, F., Crozier, S., 2009. Solving the
1068 ECG forward problem by means of standard H- and H-hierarchical adap-
1069 tive linear boundary element method: Comparison with two refinement
1070 schemes. *Biomedical Engineering, IEEE Transactions on* 56, 1454–1464.
- 1071 Surawicz, B., Childers, R., Deal, B.J., Gettes, L.S., 2009. AHA/ACCF/HRS
1072 Recommendations for the Standardization and Interpretation of the Elec-
1073 trocardiogram, Part III: Intraventricular Conduction Disturbances A Sci-
1074 entific Statement From the American Heart Association Electrocardiog-
1075 raphy and Arrhythmias Committee, Council on Clinical Cardiology; the
1076 American College of Cardiology Foundation; and the Heart Rhythm Soci-
1077 ety Endorsed by the International Society for Computerized Electrocardi-
1078 ology. *Journal of the American College of Cardiology* 53, 976–981.

- 1079 Talbot, H., Marchesseau, S., Duriez, C., Sermesant, M., Cotin, S., Delingette,
1080 H., 2013. Towards an interactive electromechanical model of the heart.
1081 *Interface Focus* 3, 20120091.
- 1082 Ten Tusscher, K., Noble, D., Noble, P., Panfilov, A., 2004. A model for human
1083 ventricular tissue. *American Journal of Physiology-Heart and Circulatory*
1084 *Physiology* 286, H1573–H1589.
- 1085 Toussaint, N., Stoeck, C.T., Sermesant, M., Schaeffter, T., Kozerke, S.,
1086 Batchelor, P.G., 2013. In-vivo human cardiac fibre architecture estimation
1087 using shape-based diffusion tensor processing. *Medical Image Analysis* 17,
1088 1243–1255.
- 1089 Van Oosterom, A., Strackee, J., 1983. The solid angle of a plane triangle.
1090 *Biomedical Engineering, IEEE Transactions on* 30, 125–126.
- 1091 Wallman, M., Smith, N.P., Rodriguez, B., 2012. A comparative study of
1092 graph-based, eikonal, and monodomain simulations for the estimation of
1093 cardiac activation times. *Biomedical Engineering, IEEE Transactions on*
1094 59, 1739–1748.
- 1095 Wallman, M., Smith, N.P., Rodriguez, B., 2014. Computational methods to
1096 reduce uncertainty in the estimation of cardiac conduction properties from
1097 electroanatomical recordings. *Medical Image Analysis* 18, 228–240.
- 1098 Wang, L., Wong, K.C., Zhang, H., Liu, H., Shi, P., 2011. Noninvasive com-
1099 putational imaging of cardiac electrophysiology for 3-d infarct. *Biomedical*
1100 *Engineering, IEEE Transactions on* 58, 1033–1043.
- 1101 Wilson, L.D., Jennings, M.M., Rosenbaum, D.S., 2011. Point: M cells are
1102 present in the ventricular myocardium. *Heart Rhythm* 8, 930–933.
- 1103 Wu, M.T., Su, M.Y.M., Huang, Y.L., Chiou, K.R., Yang, P., Pan, H.B.,
1104 Reese, T.G., Wedeen, V.J., Tseng, W.Y.I., 2009. Sequential changes of my-
1105 ocardial microstructure in patients postmyocardial infarction by diffusion-
1106 tensor cardiac MR correlation with left ventricular structure and function.
1107 *Circulation: Cardiovascular Imaging* 2, 32–40.
- 1108 Zettinig, O., Mansi, T., Georgescu, B., Kayvanpour, E., Sedaghat-Hamedani,
1109 F., Amr, A., Haas, J., Steen, H., Meder, B., Katus, H., Navab, N., Kamen,

- 1110 A., Comaniciu, D., 2013a. Fast data-driven calibration of a cardiac elec-
1111 trophysiology model from images and ECG, in: Medical Image Computing
1112 and Computer-Assisted Intervention–MICCAI 2013. Springer LNCS. vol-
1113 ume 8149, pp. 1–8.
- 1114 Zettinig, O., Mansi, T., Georgescu, B., Rapaka, S., Kamen, A., Haas, J.,
1115 Frese, K.S., Sedaghat-Hamedani, F., Kayvanpour, E., Amr, A., Hardt, S.,
1116 Mereles, D., Steen, H., Keller, A., Katus, H.A., Meder, B., Navab, N.,
1117 Comaniciu, D., 2013b. From medical images to fast computational models
1118 of heart electromechanics: An integrated framework towards clinical use,
1119 in: Functional Imaging and Modeling of the Heart. Springer LNCS. volume
1120 7945, pp. 249–258.
- 1121 Zheng, Y., Barbu, A., Georgescu, B., Scheuering, M., Comaniciu, D., 2008.
1122 Four-chamber heart modeling and automatic segmentation for 3-D cardiac
1123 CT volumes using marginal space learning and steerable features. Medical
1124 Imaging, IEEE Transactions on 27, 1668–1681.