SIAM J. SCI. COMPUT.
Vol. 44, No. 3, pp. B605–B639

PDE-AWARE DEEP LEARNING FOR INVERSE PROBLEMS IN CARDIAC ELECTROPHYSIOLOGY*

RICCARDO TENDERINI†, STEFANO PAGANI‡, ALFIO QUARTERONI§, AND SIMONE DEPARIS†

Abstract. In this work, we present a PDE-aware deep learning model for the numerical solution to the inverse problem of electrocardiography. The model both leverages data availability and exploits the knowledge of a physically based mathematical model, expressed by means of partial differential equations (PDEs), to carry out the task at hand. The goal is to estimate the epicardial potential field from measurements of the electric potential at a discrete set of points on the body surface. The employment of deep learning techniques in this context is made difficult by the low amount of clinical data at disposal, as measuring cardiac potentials requires invasive procedures. Suitably exploiting the underlying physically based mathematical model allowed circumventing the data availability issue and led to the development of fast-training and low-complexity models. Physical awareness has been pursued by means of two elements: the projection of the epicardial potential onto a space-time reduced subspace, spanned by the numerical solutions of the governing PDEs, and the inclusion of a tensorial reduced basis solver of the forward problem in the network architecture. Numerical tests have been conducted only on synthetic data, obtained via a full order model approximation of the problem at hand, and two variants of the model have been addressed. Both proved to be accurate, up to an average ℓ^1-norm relative error on epicardial activation maps of 3.5%, and both could be trained in ≈15 min. Nevertheless, some improvements, mostly concerning data generation, are necessary in order to bridge the gap with clinical applications.

Key words. cardiac electrophysiology, deep learning, ECGI, inverse problem of electrocardiography, partial differential equations, physics-informed deep learning

AMS subject classifications. 65N21, 68T05

DOI. 10.1137/21M1438529

1. Introduction.

Over the last 30 years, especially in the last decade, computational biomedical research has witnessed the development of electrocardiographic imaging (ECGI) [9, 10, 11, 27, 34, 36, 41], a novel imaging modality for noninvasive mapping of cardiac electrical activity, which makes use of body surface ECG signals and of thoracic CT-scans. The success of ECGI can be recognized in its ability of providing relevant diagnostic information on the electrical activity of the heart, visualizing it directly at the heart level and easing in turn the readability and interpretability of

*Submitted to the journal's Computational Methods in Science and Engineering section August 3, 2021; accepted for publication (in revised form) December 20, 2021; published electronically June 1, 2022.

https://doi.org/10.1137/21M1438529

Funding: SP, AQ have received funding from the *European Research Council (ERC)* under the *European Union's Horizon 2020 research and innovation programme* (grant agreement No 740132, *iHEART - An Integrated Heart Model for the simulation of the cardiac function*, P.I. Prof. A. Quarteroni). RT, SD have received funding from the *Institute of Mathematics* of EPFL and from the *Swiss National Science Foundation* (grant agreement No 200021_197021, *Data-driven approximation of haemodynamics by combined reduced order modeling and deep neural networks*).

†École Polytechnique Fédérale de Lausanne (EPFL), Institute of Mathematics, CH-1015 Lausanne, Switzerland (riccardo.tenderini@epfl.ch, simone.deparis@epfl.ch).

‡MOX, Modeling and Scientific Computing, Department of Mathematics, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133, Milano, Italy (stefano.pagani@polimi.it).

§MOX, Modeling and Scientific Computing, Department of Mathematics, Politecnico di Milano, Piazza Leonardo da Vinci 32, 10133, Milano, Italy, and École Polytechnique Fédérale de Lausanne (EPFL), Institute of Mathematics, CH-1015 Lausanne, Switzerland (Professor Emeritus) (alfio.quarteroni@polimi.it).

the results. Additionally, ECGI has proved to achieve an accuracy that plain body surface signals (as ECG or body surface potential maps) do not offer and that could be instead obtained only by means of invasive measurement techniques [27]. Despite being a recent field of research, the stunning advances made over the last decade have already allowed validation of its performance in a number of clinical practices [10, 9] and allowed it to be effectively employed in the clinical setting [35].

From a mathematical standpoint, the theoretical basis of ECGI is the so-called inverse problem of electrocardiography; it consists in finding the epicardial potential field which gives rise to body surface signals as close as possible to some target ones, in the least-squares sense. Let us consider a geometry made by the torso domain Ω_T, with the body surface Γ_B representing an external boundary and the epicardial one Γ_H representing an internal boundary. Additionally, let $\Sigma \subset \Gamma_B$ be the portion of the body surface where electric potentials are recorded and call z such recordings. Now, let $v \in H^{\frac{1}{2}}(\Gamma_H)$ be the epicardial potential and define $y(v)$ as the unique solution in $H^1(\Omega_T)$ of the following stationary elliptic problem:

$$\begin{cases} -\boldsymbol{\nabla} \cdot (D_T \boldsymbol{\nabla} y(v)) = 0 & \text{in } \Omega_T, \\ y(v) = v & \text{on } \Gamma_H, \\ \boldsymbol{\nabla} y(v) \cdot \boldsymbol{n_B} = 0 & \text{on } \Gamma_B \end{cases} \tag{1.1}$$

where D_T is the tensor expressing the electrical conductivity in the torso and $\boldsymbol{n_B}$ is the outward unit normal vector to Γ_B. Also, we define the operator $A : H^{\frac{1}{2}}(\Gamma_H) \longrightarrow H^{\frac{1}{2}}(\Sigma)$ such that

$$Av = y(v)|_\Sigma \tag{1.2}$$

and the cost functional

$$J(v) = \frac{1}{2} \int_\Sigma |y(v) - z|^2 \, d\sigma = \frac{1}{2} ||Av - z||^2_{L^2(\Sigma)} \tag{1.3}$$

for $z \in L^2(\Sigma)$. Then, the inverse problem of electrocardiography can be written as

$$\text{find } u_H \in H^{\frac{1}{2}}(\Gamma_H) : \quad J(u_H) = \min_{v \in H^{1/2}(\Gamma_H)} \left(J(v) + R(v) \right), \tag{1.4}$$

$R(v)$ being a suitable regularization term. Without stabilization (i.e., setting $R(v) = 0$) problem (1.4) is ill-posed in usual Sobolev spaces, which means that A admits an unbounded inverse operator in the spaces $H^s \ \forall s \in \mathbb{R}$ (see [48, 37, 51]). This implies that small perturbations (for instance, due to measurement errors) in the observed surface potentials z may lead to much larger variations on the reconstructed epicardial potential u_H. To retrieve well-posedness, problem (1.4) must be therefore approximated with a family of stable problems. The regularization term $R(\cdot)$ can be computed as

$$R(v) = \frac{\alpha}{2} ||v||^2_{H^{1/2}(\Gamma_H)} \quad \text{or} \quad R(v) = \frac{\alpha}{2} ||y(v)||^2_{H^1(\Omega_T)} \tag{1.5}$$

where $y(v)$ is the solution of system (1.1) with v as Dirichlet boundary datum on Γ_H and α is a hyperparameter to be tuned. Two other well-established techniques are *Tikhonov regularization* [49], which consists in computing $R(\cdot)$ as the $L^2(\Gamma_H)$-norm either of the epicardial potential or of its derivatives, and *total variation regularization*, which instead defines $R(\cdot)$ as the $L^1(\Gamma_H)$-norm of the normal derivative of the epicardial potential. Notice that all these strategies are subject to the choice of at least one hyperparameter; various techniques have been proposed to (sub-)optimize such a choice [30].

In the current work we have followed a different path, bypassing classical approaches thanks to two strategies. The first one is the exploitation of machine learning (ML) and, more specifically, of deep learning (DL), which has established itself as a pillar of a new generation of scientific development, thanks to the abundance of available data and to the progress in terms of computational power and resources. In particular, we wish to overcome the ill-posedness of the inverse problem of electrocardiography leveraging data availability, yet still somehow retaining knowledge of the physical laws governing the problem at hand. Incidentally, if until a few years ago ECGI could be regarded as almost free of ML-DL influence, recently a few works pursuing an integration between the fields have started to appear. Other than neural network (NN) models yielding heartbeats classification from body surface signals [58, 29], we ought to mention the works by Giffard-Roisin et al. [20], Wang et al. [52], Bacoyannis et al. [2], and Bujnarowski et al. [5]. In particular, [20] employs kernel ridge regression and a 2-term exponential regression to perform a patient-specific estimation of the main parameters of the chosen cardiac model. [52] builds an NN that provides (sub-)optimal values for the regularization parameters involved in the alternating direction method of multipliers iterative optimization algorithm. [2] proposes a novel method based on conditional β variational autoencoders using deep convolutional generative NNs for data-driven integration of spatio-temporal correlations and imaging information in the ECGI problem. [5] proposes a convolutional deep autoencoder which yields an imaging-free approximation of the mapping between recorded body surface potentials and epicardial ones.

The second key element is based on the observation that all classical methods to solve the inverse problem of electrocardiography involve the computation of the transfer matrix $\boldsymbol{A}$. Such a matrix depends on the geometry of the torso and of the heart and on their electrical conductivities and it can thus be inferred only through thoracic CT-scans and subsequent imaging postprocessing routines. Anyway, CT-scans imply a certain level of radiation and they are not even an available option for all patients. Developing, then, a CT-scan-free model, able to map body surface potentials to epicardial ones in a reliable and physically consistent way, could be of great help and "could serve as a preliminary study of the patient's condition, before a more thorough examination is performed" [5]. This work was developed in a *fixed geometrical setting*; with this expression, we mean that we employed the same geometry in all the numerical simulations and we kept it fixed over time, discarding variations in the shape, dimension, position, and rotation of the heart or in the dimension and shape of the torso. Furthermore, we approximated the latter as a homogeneous and isotropic volume conductor, neglecting the presence of organs other than the heart, characterized by different electric conductivities. Despite being quite strong assumptions [42, 57, 12], they can all be justified considering this work as an initial proof-of-concept on the integration between DL and ECGI. Our ultimate goal, indeed, is to work in a dynamic and parametrized geometrical setting, designing a model which does not just reconstruct the epicardial potential field, but that also provides estimates of suitable parameters characterizing the geometry and the electric conductivity.

Although [5] appears to have many points in common with this work, here a further key element is present: physical awareness. Indeed, the model we propose has been designed so that, albeit in the framework of a classical encoder-decoder structure, some elements allow us to drive the predicted solutions to be consistent with the problem physics, approximated by means of PDEs (from which the expression "PDE-aware deep learning"). Broadening the scope, the proposed model is an instance

of a physically informed DL model, i.e., a DL model which exploits the knowledge of some physical laws governing the phenomenon of interest to ease its training and/or to ameliorate its predictive power. Recent years have witnessed a flourishing of works in this direction [24, 53, 40, 32, 15, 7]; among these a special mention is deserved by Raissi, Perdikaris, and Karniadakis et al. [40], where the authors introduce physically informed neural networks, i.e., NNs able to approximate solutions to a variety of PDEs by minimizing, in the loss function, the mean squared error (MSE) on the residuals. Another work we ought to refer to is [15] by Dal Santo, Deparis, and Pegolotti; there the authors present DL models (called RB-DNNs) that, receiving in input samples of solutions to parametrized PDEs and embedding a reduced basis (RB) solver [39] as a deterministic layer inside the NN architecture, manage to reconstruct such solutions in the whole computational domain and further provide estimates of the characteristic parameters.

According to [40], physics-informed DL can perform well in the so-called small-data regime, i.e., in frameworks where the amount of data at disposal is either limited or partial or subject to a high degree of inaccuracy, because of the high cost and/or complexity of data acquisition procedures. Indeed, in such a context, classical DL methods feature severe problems in terms of robustness, generalization, and convergence, due to the lack of data. However, if the phenomenon generating such data happens to be characterized by the presence of some underlying physical laws, expressible by means of PDEs/ODEs, then classical numerical methods can be integrated with DL-based ones. In particular the knowledge of the physics of the problem at hand can be made somehow available to the DL model, acting as a physically informed regularization agent, that eases model training and improves its performances. ECGI is clearly a context of such kind; indeed on the one side recordings of the epicardial potential are nowadays possible only via intrusive techniques and assembling a dataset made of a number of observations adequate for DL applications is unfeasible. On the other side, instead, a good understanding of the most important physical laws underlying cardiac electrophysiology is present [19, 47], so that exploiting it in the framework of physically informed DL appears both doable and challenging. The goal of this work is to show how the integration between DL techniques and physically based regularization allows one to accurately solve the inverse problem of electrocardiography, even in a *small data* regime. The impact of physical awareness and of NN architecture on epicardial potentials reconstruction is investigated in various numerical test cases.

Section 2 presents a physics-informed DL model, called the space-time reduced basis deep neural network (ST-RB-DNN), that provides reliable and physically consistent solutions to the inverse problem of electrocardiography. In particular, at the beginning of the section an overview on the ST-RB-DNN model is given, highlighting its structure and its main features. Subsection 2.1 reports a brief description of the modeling and numerical approximation of cardiac electrophysiology and of the forward problem of electrocardiography. Subsection 2.2 contains an overview on the model order reduction techniques exploited in the ST-RB-DNN model. In subsection 2.3 a more extensive and technical characterization of the proposed model is reported and two different versions of it are sketched. Section 3 reports the results of the numerical tests conducted with two different versions of the ST-RB-DNN model on the same test case; a brief description of the computational setting used to generate the dataset is also given. Section 4 focuses on the discussion of the obtained results. In particular, subsection 4.1 evaluates the overall performances of the model and investigates the differences (in terms of accuracy and complexity) between its different versions, while subsection 4.2

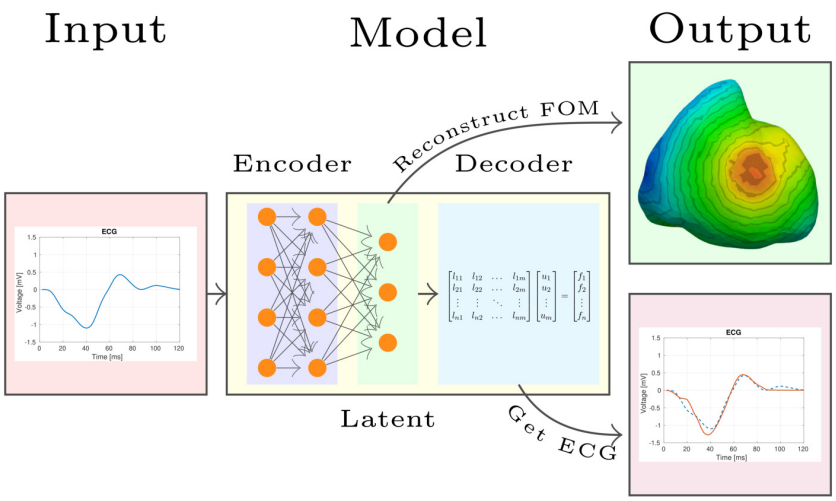

FIG. 1. *Basic structure of the ST-RB-DNN model.*

lists the main limitations and proposes some possible further developments. Section 5 provides a final summary.

2. PDE-aware deep learning models for cardiac electrophysiology. The proposed PDE-aware DL model for the inverse problem of electrocardiography is called the space-time reduced basis deep neural network, or ST-RB-DNN; its general structure is reported in the scheme of Figure 1. Three main components can be identified:

- **Input**: The input is constructed from body surface signals, recorded by electrodes placed on the patient's skin. The raw signals could possibly undergo a suitable preprocessing (lowpass/bandpass filtering, magnitude normalization, discrete Fourier transform (DFT), etc.). Additionally, all datapoints must have the same dimension, i.e., they must contain the same number of signals and all signals must have the same length.
- **Model**: The model can be identified as an autoencoder (AE), similar to the one presented in [5]. However, while the encoder consists of an NN trained with the classical backpropagation algorithm [43], the decoder is deterministic, being an embedded tensorial RB solver for problem (1.1):
  - **Encoder**: The encoder is a deep NN, which takes as input the body surface signals (or some quantities derived from them after some preprocessing routine) and learns a latent reduced representation of the epicardial potential field, responsible for the generation of the same signals given in input. To some extent, it can be stated that the encoder provides a data-driven approximation of $A^{-1} : L^2(\Sigma) \to L^2(\Gamma_H)$, i.e., the inverse of the transfer operator A defined in (1.2). Two aspects are worth considering. The first one is that the architecture of the NN depends on the nature of the quantities that are provided in input and, furthermore, it can be optimized via a grid search process, aimed at finding a (sub-)optimal balance between accuracy and complexity. The second is that the latent representation of the epicardial potential field is not uninterpretable and inexplicable, as the one proposed in [5]. Indeed, what the encoder learns are

the coefficients arising from the projection of the epicardial potential field onto a spatio-temporal reduced subspace, generated from the same physical solutions included as output in the training dataset. The way such a subspace is generated is detailed in subsections 2.1 and 2.2. We remark that body surface signals can be deterministically reconstructed from the estimated latent representation of the epicardial potential field. Moreover, such a representation features optimality properties that seamlessly depend on the method employed to generate the aforementioned reduced subspace. This entails that the epicardial potential field is encoded in the lowest possible number of coefficients that guarantee a certain accuracy level in a suitable norm.

  - **Decoder**: The decoder is responsible for the deterministic reconstruction of the body surface potentials, given the latent representation of the epicardial potential field estimated by the encoder. As such a latent representation is actually a projection onto an optimal spatio-temporal reduced subspace, the decoder can be constructed as a tensorial (to be compliant with the backpropagation training algorithm) reduced order model (ROM) solver of the forward problem of electrocardiography (FPE) [3]. In particular, our model features a decoder that employs the RB method to solve the FPE independently at all the discrete time instants. Additionally, the obtained solutions are suitably postprocessed in the full order spatial domain, so that signals matching the ones provided in input can be computed. Such a choice presents a clear advantage with respect to the usage of a trainable decoder as in [5]; indeed no additional trainable parameters are added to the ones of the encoder and the overall model complexity is potentially halved. However, designing the decoder as a ROM solver of the FPE makes it more difficult to achieve independence from imaging and CT-scans. In this paper, we have adopted a *fixed geometrical setting*, as described in the introduction; a natural future development would be to parametrize such a setting and, consequently, to adapt and update the embedded RB solver.

- **Output**: The model output is made of two parts: the latent reduced representation of the epicardial potential field and the body surface signals reconstructed by the embedded RB solver of the FPE and matching the ones given in input. Consequently, the loss functional is constructed as a weighted sum between the estimation error on the reduced epicardial potential fields $\mathcal{L}_{BC}$ and the reconstruction error on the body surface signals $\mathcal{L}_{sig}$, so that

$$\mathcal{L}(\boldsymbol{\Theta}) = \mathcal{L}_{BC}(\boldsymbol{\Theta}) + \mathcal{L}_{sig}(\boldsymbol{\Theta}) + \mathcal{L}_{reg}(\boldsymbol{\Theta}). \tag{2.1}$$

Here $\boldsymbol{\Theta}$ is the vector of the NN trainable parameters and the additional term $\mathcal{L}_{reg}$ represents a regularization term, which helps in preventing data overfitting. The precise expression of the three terms will be given in subsection 2.3.1. We remark that the ST-RB-DNN model can be trained as a pure AE, by suppressing the loss term $\mathcal{L}_{BC}$ related to the latent representation of the epicardial potential fields. In this way, the noninvasive collection of body surface potentials is sufficient for model training and all the invasively recorded cardiac potentials can be employed just for the sake of clinical validation. If we instead add the term $\mathcal{L}_{BC}$ to the loss function, also cardiac potentials are needed during the training stage. This may constitute a severe limitation when the model learns exclusively from clinical data. However, we designed the ST-RB-DNN model to encode information obtained from the numerical approximation of the FPE mathematical model. Therefore, we can include the numerical approximations of cardiac potentials collected together with the body

surface ones in the loss function, improving model performance. For these reasons, we focus mainly on the case where a penalization of the epicardial potential fields is included in the loss. Because of this, the RB solver of the FPE should be interpreted, more than as a decoder, as a physically based regularization agent, which drives the predicted epicardial potential fields to belong to a lower dimensional and physically consistent manifold. Also, we remark that the knowledge of cardiac potentials is never necessary at the testing stage, as the model input consists only of body surface signals.

Ultimately, two elements contribute to that physical awareness that makes the ST-RB-DNN model a PDE-aware DL model for ECGI:

1. The estimation of the epicardial potential field as projected onto a space-time reduced subspace, generated from the same solutions to the cardiac electrophysiology problem that represent the output part of the training dataset. In this way, indeed, the estimated epicardial potential field is forced to belong to a lower-dimensional manifold of physically consistent solutions. To meet this requirement the training dataset should be made of realistic (if not real) data, suitable for capturing the majority of the dynamics of the problem of interest.
2. The reconstruction and penalization of body surface signals equivalent to the ones provided in input, which are computed via an embedded RB solver of the FPE that acts as a deterministic decoder. Indeed, if the contribution of such a reconstruction error in the loss is sufficiently high to be nonnegligible, it drives the model to estimate epicardial potential fields which, at least in the considered *fixed geometrical setting*, determine the onset of body surface signals close to the observed ones. So, the space of admissible solutions for the inverse problem is further shrinked, as additional physical constraints have been added, and it ultimately reduces to a low-dimensional and physically consistent manifold.

As the datasets employed to train and test the proposed model have been generated *in silico*, subsection 2.1 provides an overview on the modeling and the numerical approximation of cardiac electrophysiology and of the FPE. Subsection 2.2 describes the model order reduction techniques exploited in the ST-RB-DNN model, in order to provide fast and accurate solutions to the FPE. Last, subsection 2.3 discusses in more detail the choices that have been made concerning the design and optimization of the NNs.

2.1. Data generation. Due to the paucity of clinical data relative to measurements of the electric potential at the epicardium, the datasets we employed for both training and testing the ST-RB-DNN model have been generated *in silico*. In particular, this involved the modelization and the numerical approximation of both cardiac electrophysiology and the FPE. As anticipated, all simulations have been carried out in a *fixed geometrical setting* (see section 3) and the different datapoints have been generated by randomly changing the values of the main parameters characterizing the selected cardiac model.

Cardiac electrophysiology. Cardiac electrophysiology has been modeled via the bidomain equations [50], coupled with the Aliev–Panfilov (AP) ionic model [1]. The former are characterized by a partitioning of the heart domain Ω_H in an intracellular domain Ω_H^i and an extracellular one Ω_H^e, which however coexist on the same space upon a homogenization process [21]. This is reflected in the definition of an intracellular potential u_i and of an extracellular potential u_e; also the transmembrane potential can be defined as $v := u_i - u_e$. Ultimately, the bidomain model is expressed via the following system of time-dependent PDEs:

$$
\begin{cases}
A_m\left(C_m\dfrac{\partial v}{\partial t}+I_{ion}(v,w;\boldsymbol{\mu})\right)-\boldsymbol{\nabla}\cdot(\boldsymbol{D}_i(\boldsymbol{\mu})\boldsymbol{\nabla}v)- & \\
\boldsymbol{\nabla}\cdot(\boldsymbol{D}_i(\boldsymbol{\mu})\boldsymbol{\nabla}u_e)=A_m I_{app}(\boldsymbol{\mu}) & \text{in } \Omega_H\times[t_0,T],\\
-\boldsymbol{\nabla}\cdot(\boldsymbol{D}_i(\boldsymbol{\mu})\boldsymbol{\nabla}v)-\boldsymbol{\nabla}\cdot((\boldsymbol{D}_i(\boldsymbol{\mu})+\boldsymbol{D}_e(\boldsymbol{\mu}))\boldsymbol{\nabla}u_e)=0 & \text{in } \Omega_H\times[t_0,T].
\end{cases}
\tag{2.2}
$$

Here A_m represents the area of cell membrane per unit volume, C_m the membrane capacitance per unit area, $\boldsymbol{D}_i$ and $\boldsymbol{D}_e$ the conductivity tensors in the intracellular and extracellular domains, respectively (which may take into account the anisotropy of the cardiac tissue), I_{app} an externally applied current, typically responsible for the initial activation of the heart tissue, and I_{ion} the ionic current across the cell membrane. The simulations are conducted over the time interval $[t_0;T]$. Finally, $\boldsymbol{\mu}$ represents a vector which stores all the scalar parameters that characterize the problem; notice that the expressions of I_{app}, I_{ion}, and $\boldsymbol{D}_{i,e}$ depend on $\boldsymbol{\mu}$, while the domain Ω_H is assumed to be fixed. System (2.2) must be coupled with proper initial (on v only) and boundary conditions. Concerning the latter, in particular, we have chosen to work under the *isolated heart assumption*, thus keeping the cardiac potentials independent of the one in the torso. This is achieved by imposing homogeneous Neumann boundary conditions at the epicardial surface Γ_H for both u_e and v; the impact of this simplifying assumption has been extensively analyzed and discussed, for instance, in [4]. The anisotropy of the cardiac tissue, dictated by the presence of conducting fibers, has been taken into account; specifically, different electrical conductivities have been set longitudinally and transversally to the fibers' orientation, both in the intracellular and in the extracellular domain [38]. For simplicity, we have supposed all cardiac conductivities to be constant in Ω_H (*homogeneous anisotropy assumption*) and the two transversal conductivities, coplanar and orthogonal to the fibers sheet, to be equal (*axially isotropic case*) [13]. Ultimately, the conductivity tensors could be written as

$$
\boldsymbol{D}_{i,e}(\boldsymbol{x})=\sigma_{t_{i,e}}\boldsymbol{I}+\left(\sigma_{l_{i,e}}-\sigma_{t_{i,e}}\right)\boldsymbol{a}_l(\boldsymbol{x})\boldsymbol{a}_l^T(\boldsymbol{x}),\qquad \boldsymbol{x}\in\Omega_H,
\tag{2.3}
$$

$\boldsymbol{a}_l(\boldsymbol{x})$ being a unit vector parallel to the fibers orientation at a point $\boldsymbol{x}\in\Omega_H$.

The AP model, instead, is a phenomenological ionic model that allows us to approximate the current I_{ion} induced by the movement of electrically charged ions through the cell membrane. It configures as an extension of the FitzHugh–Nagumo (FHN) model [28] and it proved to exceed its performance on cardiomyocites, leading to a much better approximation of the action potential shape and duration and of the cardiac tissue restitution curve. I_{ion} takes the form of a cubic nonlinear function of v and a single (dimensionless) gating variable w plays the role of a recovery function, allowing one to model the refractoriness of cells disregarding any subcellular process. From a mathematical standpoint, the AP model is expressed by the following equations:

$$
\begin{cases}
I_{ion}(v,w;\boldsymbol{\mu})=Kv(v-a)(v-1)+v & \textit{in } \Omega_H\times[t_0,T],\\
\dfrac{\partial w}{\partial t}=g(v,w;\boldsymbol{\mu})=C(w)(-w-Kv(v-b-1)) & \textit{in } \Omega_H\times[t_0,T],
\end{cases}
\tag{2.4}
$$

where the parameters $K,a,b,\epsilon_0,\mu_1,\mu_2$ are all related to the cell. In particular, a acts as an oscillation threshold, above which an action potential is fired, and the weighting term

$$
C(w)=\epsilon_0+\frac{\mu_1 w}{v+\mu_2}
\tag{2.5}
$$

has been specifically added (compared to the FHN model) to account for a finer tuning of the myocardial restitution curve. Additionally, a suitable initial condition has to

be imposed on w (typically $w(t = t_0) = 0$) and, in order to get the desired scaling properties, it is necessary to scale the time variable as $t[ms] = 12.9\ t[t.u.]$.

Upon a Galerkin finite element (FE) approximation, the discrete-in-space continuous-in-time formulation of the bidomain equations coupled with any phenomenological ionic model reads as follows.

PROBLEM 1. *Given* $\boldsymbol{\mu} \in \mathcal{P} \subset \mathbb{R}^p$, *find* $\boldsymbol{u}_{e_h} = \boldsymbol{u}_{e_h}(t;\boldsymbol{\mu})$, $\boldsymbol{v}_h = \boldsymbol{v}_h(t;\boldsymbol{\mu})$, *and* $\boldsymbol{w}_h = \boldsymbol{w}_h(t;\boldsymbol{\mu})$ *such that*

$$
\begin{cases}
A_m C_m \boldsymbol{M}\dfrac{\partial \boldsymbol{v}_h}{\partial t} + \boldsymbol{A}_{in}(\boldsymbol{\mu})\boldsymbol{v}_h + \boldsymbol{A}_{in}(\boldsymbol{\mu})\boldsymbol{u}_{e_h} & t \in [t_0, T], \\
+\ A_m \boldsymbol{I}_{ion}(\boldsymbol{v}_h, \boldsymbol{w}_h; \boldsymbol{\mu}) = A_m \boldsymbol{I}_{app}(t;\boldsymbol{\mu}), & \\
\boldsymbol{A}_{in}(\boldsymbol{\mu})\boldsymbol{v}_h + \big(\boldsymbol{A}_{in}(\boldsymbol{\mu}) + \boldsymbol{A}_{ex}(\boldsymbol{\mu})\big)\boldsymbol{u}_{e_h} = 0, & t \in [t_0, T], \\
\dfrac{\partial \boldsymbol{w}_h}{\partial t} = g(\boldsymbol{v}_h, \boldsymbol{w}_h; \boldsymbol{\mu}), & t \in [t_0, T], \\
\boldsymbol{v}_h(t_0;\boldsymbol{\mu}) = \boldsymbol{v}_0(\boldsymbol{\mu}); \qquad \boldsymbol{w}_h(t_0;\boldsymbol{\mu}) = \boldsymbol{w}_0(\boldsymbol{\mu}), &
\end{cases}
\tag{2.6}
$$

where $\boldsymbol{M}$ *is the mass matrix and* $\boldsymbol{A}_{in}$, $\boldsymbol{A}_{ex}$ *are the intracellular and extracellular stiffness matrices, respectively.*

The forward problem of electrocardiography. In modeling the FPE, we can approximate the human torso as an isotropic and inhomogeneous volume conductor; indeed, as reported in [47], the torso is not made of excitable cells, as the heart, and thus there is no need for a model able to simulate the behavior of the transmembrane potential. The inhomogeneity derives from the presence of different organs (bones, lungs, blood vessels, etc.), which are characterized by different electric conductivities.

Under the *isolated heart assumption*, at any $t \in [t_0; T]$ the torso potential u_T can be computed by solving a generalized Laplace equation of the form

$$
-\boldsymbol{\nabla} \cdot (\boldsymbol{D}_T \boldsymbol{\nabla} u_T) = 0 \qquad \text{in } \Omega_T \tag{2.7}
$$

coupled with a homogeneous Neumann boundary condition at the body surface Γ_B and with an inhomogeneous Dirichlet one at the epicardium Γ_H, where the extracellular potential u_e acts as boundary datum. In (2.7) Ω_T is the reference torso domain and $\boldsymbol{D}_T$ represents the torso conductivity tensor.

As discussed in [4], working under the *isolated heart assumption* provides the advantage of getting a *one-way coupling* between the heart and torso problems, easing and speeding up the computations. It also prevents imposing the continuity of the electric potential fluxes at the heart-torso interface; in terms of body surface potentials, this results in signals that often show the correct shape but that feature abnormal magnitudes.

Upon a Galerkin FE approximation, the problem reads as follows.

PROBLEM 2. *Given* $\boldsymbol{\mu} \in \mathcal{P} \subset \mathbb{R}^p$, *find* $\boldsymbol{u}_{t_h} = \boldsymbol{u}_{t_h}(t;\boldsymbol{\mu})$ *such that*

$$
\boldsymbol{A}(\boldsymbol{\mu})\boldsymbol{u}_{t_h}(t) = -\boldsymbol{A}_{\Gamma_H}(\boldsymbol{\mu})\boldsymbol{u}_{e_h}^{\Gamma_H}(t), \qquad t \in [t_0, T], \tag{2.8}
$$

where $\boldsymbol{A}_{\Gamma_H}(\boldsymbol{\mu}) =: \boldsymbol{A}(\boldsymbol{\mu})[\cdot, \{j_{Dir}\}]$, *where* $\boldsymbol{A}$ *is the torso stiffness matrix and* $\{j_{Dir}\}$ *the set of degrees of freedom (DOFs) at which the Dirichlet boundary condition is imposed. Furthermore* $\boldsymbol{u}_{e_h}^{\Gamma_H}(t)$ *represents the FE discretization of the trace of the extracellular potential* u_e *on the epicardial surface* Γ_H *at the time instant* t.

2.2. Model order reduction. In the ST-RB-DNN model, ROM techniques have been used both to estimate physically meaningful and reduced latent representations of the epicardial potential fields and to efficiently solve the generalized Laplace equation in the torso, in order to reconstruct signals matching the ones given in input. Considering that Problem 2 is stationary, the latter task can be carried out by resorting to the classical RB method [39], i.e., by deriving a basis for a reduced subspace in space (via proper orthogonal decomposition (POD) of the snapshots' tensor), performing a projection along the spatial dimension, and solving the resulting reduced problem independently at all the discrete time instants. The former task, instead, can be much better accomplished by resorting to spatio-temporal ROM techniques that allow one to encode the information coming from time-dependent fields into a very low number of coefficients, (almost) independent of the level of refinement of the grid along both the spatial and the temporal dimension. In particular, the space-time-reduced subspace where the epicardial potential fields have been projected has been generated using the *tailored temporal subspaces via the ST-HOSVD* approach proposed in [8]. This strategy is based on the computation of a tailored temporal basis for each element of the spatial one, allowing for a dramatic reduction of the dimensionality of the resulting spatio-temporal subspace. We notice that estimating a full order model (FOM) approximation of the epicardial potential field is not a viable option in the framework of the ST-RB-DNN model as that would significantly increase the overall complexity, severely hindering the training process. Reference [31] features a detailed theoretical analysis on this aspect, which can be also adapted to the present case.

In the following, we employ the aforementioned ROM techniques in the ST-RB-DNN model, in order to efficiently solve the FPE inside an embedded tensorial solver. Consider the third-order tensor $\mathcal{X} \in \mathbb{R}^{N_h \times N_t \times N_{\boldsymbol{\mu}}}$ storing the solutions to (2.8) at all the space-time $N_h N_t$ DOFs and for $N_{\boldsymbol{\mu}}$ different parameter values. The basis in space, encoded by the matrix $\boldsymbol{V}_s \in \mathbb{R}^{N_h \times n_h}$, is computed by applying a truncated POD to the mode-1 unfolding of $\mathcal{X}$. Once such a basis is derived, by projecting all quantities onto the reduced subspace it spans, we get

$$\tilde{\boldsymbol{A}}(\boldsymbol{\mu})\tilde{\boldsymbol{u}}_{t_h}^{(l)} = \tilde{\boldsymbol{u}}_{e_h}^{\Gamma_H{}^{(l)}}(\boldsymbol{\mu}), \qquad l \in \{0, \dots, N_t\}, \tag{2.9}$$

where, with $\boldsymbol{A}(\boldsymbol{\mu})$, $\boldsymbol{A}_{\Gamma_H}(\boldsymbol{\mu})$, and $\boldsymbol{u}_{e_h}^{\Gamma_H}(\boldsymbol{\mu})$ defined as in Problem 2,

$$\tilde{\boldsymbol{A}}(\boldsymbol{\mu}) =: \boldsymbol{V}_s^T \boldsymbol{A}(\boldsymbol{\mu}) \boldsymbol{V}_s \qquad \in \mathbb{R}^{n_h \times n_h}, \tag{2.10a}$$

$$\tilde{\boldsymbol{u}}_{t_h}^{(l)} =: \boldsymbol{V}_s^T \boldsymbol{u}_{t_h}^{(l)} \qquad \in \mathbb{R}^{n_h}, \quad l \in \{0, \dots, N_t\}, \tag{2.10b}$$

$$\tilde{\boldsymbol{u}}_{e_h}^{\Gamma_H{}^{(l)}}(\boldsymbol{\mu}) =: -\boldsymbol{V}_s^T \boldsymbol{A}_{\Gamma_H}(\boldsymbol{\mu}) \boldsymbol{u}_{e_h}^{\Gamma_H{}^{(l)}}(\boldsymbol{\mu}) \qquad \in \mathbb{R}^{n_h}, \quad l \in \{0, \dots, N_t\}. \tag{2.10c}$$

Based on (2.8), we can perform a dimensionality reduction step also on the epicardial extracellular potential. So, let $\mathcal{X}^e \in \mathbb{R}^{N_h^e \times N_t \times N_{\boldsymbol{\mu}}}$ be the third-order tensor storing the values of the epicardial extracellular potential $\boldsymbol{u}_{e_h}^{\Gamma_H}$ at all the $N_h^e N_t$ space-time DOFs of the epicardial surface and for $N_{\boldsymbol{\mu}}$ different parameter values. Then an RB in space, stored by the matrix $\boldsymbol{V}_s^e \in \mathbb{R}^{N_h^e \times n_h^e}$, can be computed by applying a truncated POD to the mode-1 unfolding of $\mathcal{X}^e$. Equation (2.8) can be written in reduced form as

$$\tilde{\boldsymbol{A}}(\boldsymbol{\mu})\tilde{\boldsymbol{u}}_{t_h}^{(l)} = \tilde{\boldsymbol{A}}^e(\boldsymbol{\mu})\tilde{\boldsymbol{u}}_{e_h}^{\Gamma_H{}^{(l)}}(\boldsymbol{\mu}), \qquad l \in \{0, \dots, N_t\}, \tag{2.11}$$

where $\tilde{\boldsymbol{A}}(\boldsymbol{\mu})$ and $\tilde{\boldsymbol{u}}_{t_h}^{(l)}$ are defined as in (2.10a) and (2.10b), respectively, while

$$\tilde{\boldsymbol{A}}^e(\boldsymbol{\mu}) =: \boldsymbol{V}_s^T \boldsymbol{A}_{\Gamma_H}(\boldsymbol{\mu}) \boldsymbol{V}_s^e \qquad \in \mathbb{R}^{n_h \times n_h^e}, \tag{2.12a}$$

$$\tilde{\boldsymbol{u}}_{e_h}^{\Gamma_H{}^{(l)}}(\boldsymbol{\mu}) =: -\boldsymbol{V}_s^{e^T} \boldsymbol{u}_{e_h}^{\Gamma_H{}^{(l)}}(\boldsymbol{\mu}) \qquad \in \mathbb{R}^{n_h^e}, \qquad l \in \{0, \dots, N_t\}. \tag{2.12b}$$

As discussed before, in the framework of the ST-RB-DNN model the epicardial potential field is projected also along the temporal dimension, as it is extremely convenient to encode its dynamics in a low and grid-refinement-independent number of coefficients. Following the *tailored temporal subspaces via the ST-HOSVD* approach, n_h^e different truncated POD are applied to the projections of the snapshots' tensor $\mathcal{X}^e$ onto the spaces spanned by the different spatial basis elements, i.e.,

$$\mathcal{X}^e(\boldsymbol{V}_{s_i}^e) \in \mathbb{R}^{N_t \times N_{\boldsymbol{\mu}}} \quad s.t. \quad \left(\mathcal{X}^e(\boldsymbol{V}_{s_i}^e)\right)_{jk} = \sum_{l=1}^{N_h^e} \mathcal{X}_{ljk}^e \boldsymbol{V}_{s_{il}}^e, \quad i \in \{1, \dots, n_h^e\}. \tag{2.13}$$

These POD allow us to compute the temporal bases $\boldsymbol{V}_{t_i}^e \in \mathbb{R}^{N_t \times n_t^i}\ \forall\ i \in \{1, \dots, n_h^e\}$. The overall space-time RB for the epicardial potential is then made of $n_{st} =: \sum_{i=1}^{n_h^e} n_t^i$ elements of dimension $N_h^e \times N_t$, each one being defined as the outer product between an element of the RB in space $\boldsymbol{V}_s^e$ and an element of the associated RB in time $\{\boldsymbol{V}_{t_i}^e\}_{i=1}^{n_h^e}$.

The embedded reduced solver acting as decoder in the ST-RB-DNN model computes the ROM-in-space FOM-in-time solution to the generalized Laplace equation in the torso (see (2.11)) receiving as input the space-time projection of the epicardial potential field. Leveraging orthonormality due to the POD, such a projection can be expressed in terms of the FOM approximation as

$$\left(\hat{\boldsymbol{u}}_{e_h}^{\Gamma_H}(\boldsymbol{\mu})\right)_k = \sum_{i=0}^{N_h^e} \sum_{j=0}^{N_t} \boldsymbol{\pi}_{ij}^k \left(\boldsymbol{u}_{e_h}^{\Gamma_H}(\boldsymbol{\mu})\right)_{ij}, \qquad k \in \mathbb{N}(n_{st}), \tag{2.14}$$

where $\boldsymbol{\pi}_{ij}^k$ denotes the element in position (i, j) of the kth space-time basis function. Therefore, the ROM-in-space FOM-in-time epicardial potential $\tilde{\boldsymbol{u}}_{e_h}^{\Gamma_H}(\boldsymbol{\mu})$ can be derived by expanding in time $\hat{\boldsymbol{u}}_{e_h}^{\Gamma_H}(\boldsymbol{\mu})$ as

$$\left(\tilde{\boldsymbol{u}}^{\Gamma_H{}^{(l)}}(\boldsymbol{\mu})\right)_i = \sum_{j=0}^{n_t^i} \left(\boldsymbol{V}_{t_i}^e\right)_{jl} \hat{\boldsymbol{u}}_{\mathcal{F}(i,j)}^{\Gamma_H}(\boldsymbol{\mu}), \quad i \in \{1, \dots, n_h^e\}\ \ l \in \{0, \dots, N_t\}, \tag{2.15}$$

where

$$\mathcal{F} : \left(i \in \mathbb{N}(n_h^e),\ j \in \mathbb{N}(n_t^i)\right) \to \sum_{k=1}^{i-1} n_t^k + j \quad \in \mathbb{N}(n_{st}) \tag{2.16}$$

is the mapping from the space and time bases indexes to the space-time basis index.

Remark. If the torso is approximated as an inhomogeneous volume conductor, featuring the presence of organs with different electric conductivities, then it is possible to take advantage of an affine parametrization to efficiently assemble the stiffness matrix during the online stage. For instance, if we suppose the torso to be partitioned into N_p parts with conductivities $\boldsymbol{\mu} =: \{\sigma_p\}_{p=1}^{N_p}$, then the stiffness matrix can be written as

$$\boldsymbol{A}(\boldsymbol{\mu}) = \sum_{p=1}^{N_p} \sigma_p \boldsymbol{A}_p, \tag{2.17}$$

where $\boldsymbol{A}_p$ is the stiffness matrix associated to the DOFs belonging to the part p of the torso and computed assuming a default conductivity of 1. Readily, an equivalent affine decomposition applies also to the reduced torso stiffness matrix $\tilde{\boldsymbol{A}}(\boldsymbol{\mu})$.

2.3. Neural networks architectures. This subsection is devoted to a more detailed presentation of the ST-RB-DNN model architecture. In particular, the first part provides a description of the generic model structure (i.e., the general structure of the ST-RB-DNN model, independently of the nature of the signals given in input). The second and the third part, instead, focus on the two models that have been actually implemented and tested in this work, analyzing their specific features with respect to the general case.

2.3.1. The generic ST-RB-DNN model. The architecture of the generic ST-RB-DNN model is reported in Figure 2. The model input (in red) consists of body surface signals or of quantities that can be directly derived from the latter. The initial processing is done by a trainable NN (in blue) whose actual architecture depends on the nature of the quantities given in input and which acts as an encoder. Such NN is responsible for the estimation of two type of quantities:

1. *A latent representation of the epicardial potential field* (in orange): The main element which is estimated by the encoder is a latent representation of the epicardial potential field, which is obtained by means of a projection onto a space-time reduced subspace generated from physical solutions to the heart electrophysiology problem (see (2.14)). Notice that the latent representation estimated by the encoder, called $\bar{\boldsymbol{u}}_{e_h}^{a,\boldsymbol{\mu}}$, is not the actual projection to be lately provided in input to the embedded RB solver of the FPE (in cyan), called $\hat{\boldsymbol{u}}_{e_h}^{a,\boldsymbol{\mu}}$. Indeed, the former is processed by a deterministic *rescaler layer* (in yellow) to obtain the latter; the rescaling is affine and it has the following form:

$$\hat{\boldsymbol{u}}_{e_h}^{a,\boldsymbol{\mu}} = \hat{\boldsymbol{u}}_{e_h}^{shift} + \hat{\boldsymbol{u}}_{e_h}^{scale} \bar{\boldsymbol{u}}_{e_h}^{a,\boldsymbol{\mu}} \tag{2.18}$$

with $\hat{\boldsymbol{u}}_{e_h}^{shift}$ and $\hat{\boldsymbol{u}}_{e_h}^{scale}$ representing the additive and the multiplicative terms of the affine transformation. In simpler terms, the NN estimates how far (in terms of the epicardial potential) the current datapoint is from a target value, referred to as the "shift," with a scaling factor given by what is referred to as the "scale." Several choices can be made for the values of these two quantities; we decided to

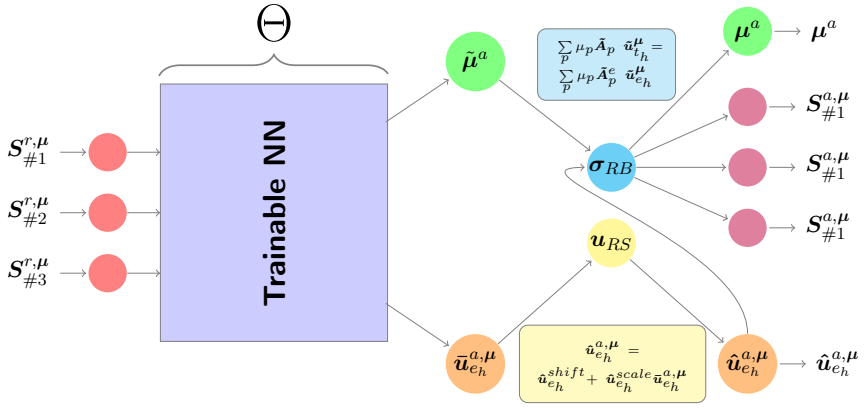

FIG. 2. *Scheme of the generic architecture of the ST-RB-DNN model.*

set $\hat{\boldsymbol{u}}_{e_h}^{shift}$ to $\mathbf{0}$ (so, corresponding to a totally inactivated heart) and to compute $\hat{\boldsymbol{u}}_{e_h}^{scale}$ as the standard deviation of the epicardial potentials in the training dataset, assuming them to have $\mathbf{0}$ mean. The rescaling is motivated by the fact that the values of the coefficients arising from the projection onto the space-time-reduced subspace happen to span over a broad range of orders of magnitude and this severely hinders the training process. Exploiting (2.18) significantly improves the performance.

2. *Torso conductivities* (in green): Other than the epicardial potential field, the encoder may also learn the values of the electric conductivities of the different organs considered in the torso geometry, in case an inhomogeneous approximation of the latter is performed. To be precise, following the suggestions given in [15], such parameters are estimated as normalized in the interval $[0; 1]$ by employing a *sigmoid* as an activation function. Additionally, working under the *isolated heart assumption*, what matters are the relative values of the conductivities and not their absolute ones; thus, one conductivity must always be normalized to 1 by default and so no parameters' estimation is needed if a homogeneous torso approximation is employed.

Finally, the quantities estimated by the encoder are fed to an embedded tensorial RB solver of the FPE (in cyan), which acts as an efficient and deterministic decoder:

1. At each forward pass during the training phase, the decoder takes as input the normalized values of the torso conductivities and the estimated rescaled space-time-reduced epicardial potential $\hat{\boldsymbol{u}}_{e_h}^{a,\boldsymbol{\mu}} \in \mathbb{R}^{b_s \times n_{st} \times 1}$. Here b_s denotes the batch size used by the stochastic gradient descent algorithm.
2. The decoder assembles the left-hand side and right-hand side stiffness matrices appearing in (2.11), eventually exploiting affine parametrization (see (2.17)).
3. The decoder expands the space-time-reduced epicardial potentials $\hat{\boldsymbol{u}}_{e_h}^{a,\boldsymbol{\mu}} \in \mathbb{R}^{b_s \times n_{st} \times 1}$ along the temporal dimension, thus computing $\tilde{\boldsymbol{u}}_{e_h}^{a,\boldsymbol{\mu}}$.
4. The decoder solves the ROM FPE (2.11) so that the ROM-in-space FOM-in-time torso potential $\tilde{\boldsymbol{u}}_{t_h}^{a,\boldsymbol{\mu}} \in \mathbb{R}^{b_s \times n_h \times N_t}$ is computed.
5. Finally, the decoder computes signals that are analogous to the ones provided in input to the model. The way this task is carried out depends on the nature of the input signals.

As reported in (2.1), the loss functional is constructed as the sum of three contributions: the error on the reduced epicardial potential field, the error on the reconstruction of body surface signals, and a *ridge regularization* term, which helps in preventing *overfitting*. In more detail, the loss functional is defined as follows:

$$\begin{aligned}(2.19)\quad \mathcal{L}(\boldsymbol{\Theta}) &= \mathcal{L}_{BC}(\boldsymbol{\Theta}) + \mathcal{L}_{sig}(\boldsymbol{\Theta}) + \mathcal{L}_{reg}(\boldsymbol{\Theta}) \\ &= w_{BC}\ MAE_{\boldsymbol{\sigma}}\big(\hat{\boldsymbol{u}}_{e_h}^{r}, \hat{\boldsymbol{u}}_{e_h}^{a}(\boldsymbol{\Theta})\big)\ + \sum_{i=1}^{n_{sig}} MSE\big(\boldsymbol{S}_{\#i}^{r}, \boldsymbol{S}_{\#i}^{a}(\boldsymbol{\Theta})\big) +\ \lambda_r ||\boldsymbol{\Theta}||_2^2\ .\end{aligned}$$

The quantities with apex ^{r} denote target values, the quantities with apex ^{a} denote approximated values, $\boldsymbol{S}_{\#i}$ encodes the i^{th} input signal, from a total of n_{sig} signals, and $w_{BC} \in \mathbb{R}$ represents the weight of the space-time-reduced epicardial potential field error. The choice of using the Mean Absolute Error (MAE) instead of the Mean Squared Error (MSE) for the loss term $\mathcal{L}_{BC}$ aims at forcing to 0 the least relevant coefficients, as characteristic of ℓ^1-norm penalization. Additionally, the MAE appearing in $\mathcal{L}_{BC}$ is weighted not only by the nonnegative scalar w_{BC}, but also by the singular

values that, during the POD construction of the two RBs, have been associated to the different coefficients. So,

$$MAE_{\boldsymbol{\sigma}}\left(\hat{\boldsymbol{u}}^r_{e_h}, \hat{\boldsymbol{u}}^a_{e_h}\right) =: \frac{1}{b_s n_{st}} \sum_{i=1}^{b_s} \sum_{j=1}^{n_h^e} \sum_{k=1}^{n_t^j} \sqrt{\frac{\sigma_j^{e,s}\sigma_k^{e,t,j}}{\sigma_1^{e,s}\sigma_1^{e,t,1}}} \left| \left(\hat{\boldsymbol{u}}^r_{e_h} - \hat{\boldsymbol{u}}^a_{e_h}\right)_{\mathcal{F}(j,k)} \right|, \tag{2.20}$$

where $\{\sigma_j^{e,s}\}_{j=1}^{n_h^e}$ are the singular values arising from the (spatial) POD applied to the mode-1 unfolding of the epicardial potentials tensor, $\{\sigma_k^{e,t,j}\}_{k=1}^{n_t^j}$ are the singular values deriving from the (temporal) POD applied to the projection of the epicardial potentials tensor onto the one-dimensional (1D) subspace spanned by the jth spatial basis element, b_s is the training batch size, and $\mathcal{F}(\cdot,\cdot)$ is the mapping from the space and time bases indexes to the space-time basis index, as defined in (2.16). In this way, the weight of the error on a space-time-reduced epicardial potential coefficient in the loss gets lower and lower as its relevance decreases; this allows us to obtain better estimates of the most relevant coefficients, easing the training process and improving model performances.

Three remarks follow. First, by resorting to ROM techniques, we managed to mitigate the effects of the ill-posedness of the inverse problem, as the dimensionality of the solution space is dramatically reduced. However, instabilities might arise when the number of considered spatio-temporal basis functions (i.e., as the POD tolerances decrease). Due to the small dimension of our dataset, we did not experience any problem in this sense; nevertheless, adding supplementary regularization terms—as the ones of (1.5)—to the loss functional ensures well-posedness. Second, when training the ST-RB-DNN model as a pure AE, we indeed considered a loss regularization term $\mathcal{L}_{reg}^{AE}$ of the form

$$\mathcal{L}_{reg}^{AE} = \lambda_r ||\boldsymbol{\Theta}||_2^2 + \lambda_t ||\boldsymbol{u}^r_{t_h}||^2_{H^1(\Omega_T)} + \lambda_e ||\boldsymbol{u}^r_{e_h}||^2_{L^2(\Gamma_H)}. \tag{2.21}$$

The values of the hyperparameters λ_t and λ_e have been chosen according to the results of a grid search process. Third, the estimation error on the torso conductivities is not involved in the expression of the loss functional; as in [15], such values are obtained as a by-product.

2.3.2. The time-series-based ST-RB-DNN model. The architecture of the time-series-based ST-RB-DNN model is shown in Figure 3. Its main feature, with respect to the generic model, is that it takes as input body surface signals organized in the form of time series. The following aspects deserve to be discussed:

- As suggested in [58, 29], at the preprocessing stage it is useful to remove noise components from body surface signals, since it eases the training process and it increases the representative power of the models. Many works regarding the denoising of real body surface potentials are available, such as [44, 45]. In this project we considered only numerically generated signals with superimposed white Gaussian noise, proportional to the signals themselves at each discrete time instant and at an average signal-to-noise ratio (SNR) of 26 dB; thus, we employed a standard low-pass Butterworth filter of order $n = 3$ [6]. Additionally, all signals have been normalized in the interval $[-1; 1]$ to ensure they are all within the same scale, as in [5].
- Since the input signals are provided as time series, we chose to employ a convolutional neural network (CNN) design for the encoder. In particular, the structure of the CNN can be divided in to three parts:

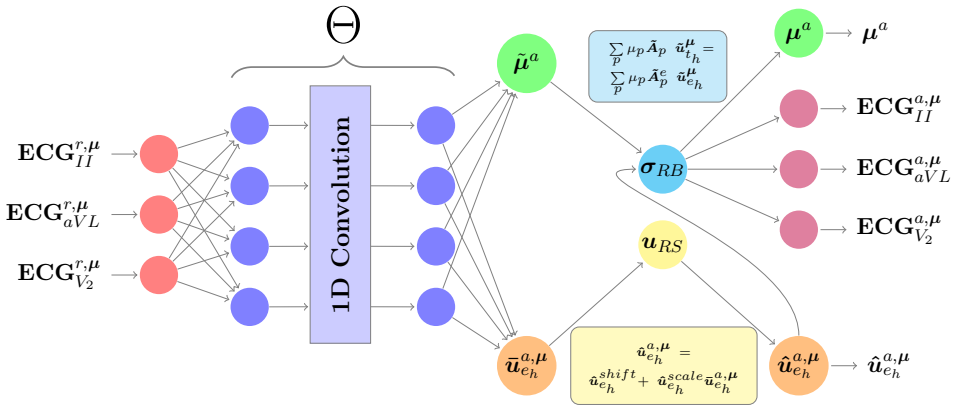

FIG. 3. *Scheme of the architecture of the time-series-based ST-RB-DNN model. Notice that as the input body surface signals are in the form of time series, the trainable NN acting as decoder is a CNN, with* 1*D convolutional layers acting along the temporal dimension only.*

- *Preconvolutional fully connected layers*: As the first step, the input signals are processed by some fully connected layers; their aim is to extract features on top of which convolutional operations may exhibit better performance. It is relevant to notice that these layers only combine quantities related to the same time instant, using the same set of weights at all time instants. In other terms, the original input signals are nonlinearly combined together, so that other signals, featuring a better information encoding, can be derived. The activation function is rectified linear unit (ReLU) for all neurons.
- 1*D convolutional + Max-Pooling layers*: After the initial "preprocessing" stage, the resulting data are passed by to 1D convolutional layers that convolve them along the temporal dimension. The number of layers, the dimensionality of the convolutional kernel, and the number of convolutional filters are three key hyperparameters of the model. Each 1D-convolutional layer is followed by a *Max-Pooling* layer, with pooling window of dimension 2×3 and with 2×2 stride. As a result, with proper padding choices, the size of the input is halved along both dimensions at each step, reducing in turn the number of trainable parameters of the model.
- *Postconvolutional fully connected layers*: After 1D convolutions, the data are finally processed by a second set of fully connected layers. Their aim is to extract relevant features to ultimately perform a good estimation of both the torso conductivities and the space-time-reduced epicardial potential. All neurons of all layers are ReLU-activated, except from the ones of the last layer, i.e., the one providing the final estimates. These indeed feature two different activation functions: *sigmoid* for the neurons storing the normalized values of the torso conductivities and *SELU* (i.e., an optimally scaled version of the exponential linear unit) for the ones storing the space-time-reduced epicardial potential. This last choice allows one to circumvent the so-called dying ReLU problem [33], which has been experienced in some preliminary tests employing ReLU as an activation function.

Additionally, *dropout* layers with $p = 0.20$ are added to the model to reduce the risk of *overfitting*.

2.3.3. The DFT-based ST-RB-DNN model. The architecture of the DFT-based ST-RB-DNN model can be visualized in Figure 4. The name stems from the

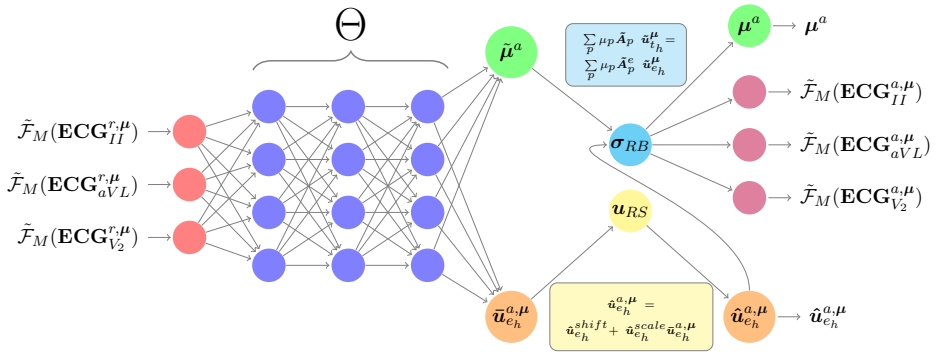

FIG. 4. *Scheme of the architecture of the DFT-based ST-RB-DNN model. Notice that as the body surface signals are given in input by means of their lowest-frequency DFT coefficients, the trainable NN acting as decoder is an MLP, since temporal dynamics have already been handled by the DFT itself at the preprocessing stage.*

usage of low-frequency DFT coefficients of the body surface signals to assemble the model input. The 1D DFT is the discrete counterpart of the univariate Fourier transform and it turns a sequence of N complex numbers $\{x_n\}_{n=0}^{N-1}$, which belongs to the temporal domain, into another sequence of N complex numbers $\{X_k\}_{k=0}^{N-1}$, which belongs instead to the frequency domain. In particular,

$$X_k = \sum_{n=0}^{N-1} x_n \, e^{-\frac{2\pi i}{N} kn} = \sum_{n=0}^{N-1} x_n \left[\cos\left(\frac{2\pi}{N} kn\right) - i \sin\left(\frac{2\pi}{N} kn\right)\right]. \tag{2.22}$$

As (2.22) is hermitian on real inputs, the dimensionality of the DFT applied to body surface signals is $\bar{M} =: \lfloor N_t/2 \rfloor + 1$, where N_t is the signal length. The dimensionality of the model inputs is then $M \leq \bar{M}$, upon having excluded a certain portion of coefficients, related to the highest-frequency modes; the choice of M is a model hyperparameter. The DFT of body surface signals have been efficiently computed using the fast Fourier transform (FFT) algorithm [14]. Additionally, a bisymmetric logarithmic transform of the form

$$\hat{x} = sign(x) \log_{10}\left(1 + \left|\frac{x}{C}\right|\right) \tag{2.23}$$

with $C = 1/\ln(10)$ [54] has been applied to both the real and the imaginary parts of the selected DFT coefficients to narrow the width of their orders of magnitude. We underline the following aspects:

- A drawback of the time-series-based ST-RB-DNN model is that it can work only at a fixed acquisition frequency; in our test cases, for instance, we sampled body surface signals at 500 Hz. Thus, if the signals to be processed have been sampled at a lower frequency, they have to be interpolated in time before being provided to the model. Conversely, the lowest-frequency DFT coefficients feature a low sensitivity with respect to the acquisition frequency; thus, using them as input allows us to avoid any interpolation at the preprocessing stage. Notice, as a caveat, that the low sensitivity occurs only with respect to the lowest-frequency DFT coefficients; however highest-frequency ones are typically not very informative, so that model accuracy should not be too much affected by their removal.
- Since the application of a DFT at the preprocessing stage already allows us to capture temporal dynamics, a convolution inside the encoder is no longer needed.

Thus, the latter simply consists of a multiple layer perceptron (MLP), with several flanked fully connected layers. Also, the encoder is equipped with *dropout* layers with $p = 0.80$ (i.e., 20% of the neurons are dropped at each update of the training phase) to reduce the risk of *overfitting* [46] and all its layers are ReLU-activated, except the last one. Its neurons, indeed, feature either sigmoid or SELU as the activation function, depending on the nature of the quantity they are estimating.

3. Results. In this section, we present the numerical results got with the ST-RB-DNN model on a benchmark test case. The data used to train and test the developed models have been generated numerically. In particular, heart electrophysiology has been approximated by solving the bidomain equations coupled with the phenomenological AP ionic model, on a reference biventricular geometry (taken from [26]—see Figure 5). The FPE required instead to solve a generalized Laplace equation in the human torso. Again, we considered a fixed geometry (taken from [18]—see Figure 6); moreover, we modeled the torso as a homogeneous and isotropic volume conductor, discarding the presence of different organs (as bones, lungs, or blood vessels) with different electric conductivities. Working under the *isolated heart assumption*, the two problems could be fully decoupled and solved sequentially, one after the other. All numerical simulations have been carried out employing the classical FE method, with elements

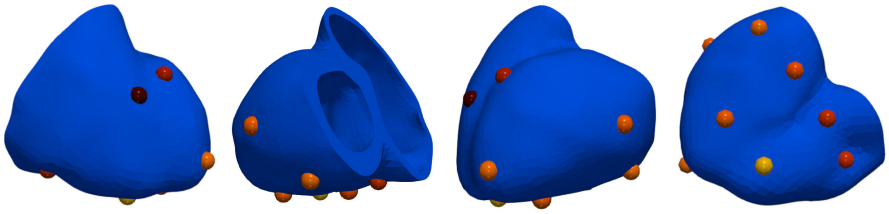

FIG. 5. *Geometry of the biventricle employed in the FOM simulations, taken from* [26]. *The colored spheres denote the sites of potential occurrence of epicardial breakthroughs in a healthy patient, according to* [56]*; the colors of the spheres are associated to the probability of epicardial breakthrough occurrence (refer to the colormap on the left side). Left: anterior view. Center-left: posterior view. Center-right: left lateral view. Right: inferior view.*

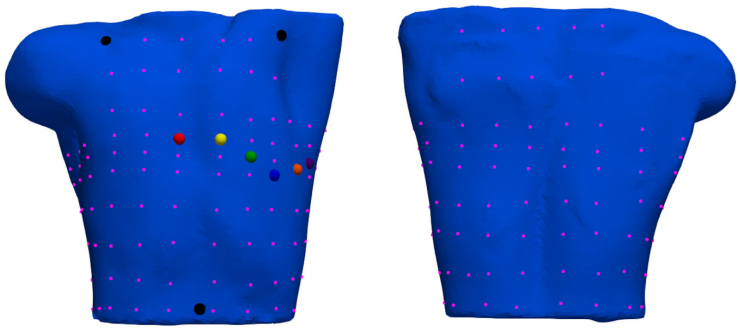

FIG. 6. *Geometry of the human torso employed in the FOM simulations, taken from* [18]. *The black spheres denote the positions of the* 3 *peripheral electrodes used in the standard* 12*-lead ECG system. The colored spheres denote the positions of the* 6 *precordial electrodes; their colors have been set according to the American Heart Association color-coding system. The lilac squares, together with the colored spheres, denote the positions of the electrodes on the "simulated" vest. Left: anterior view. Right: posterior view.*

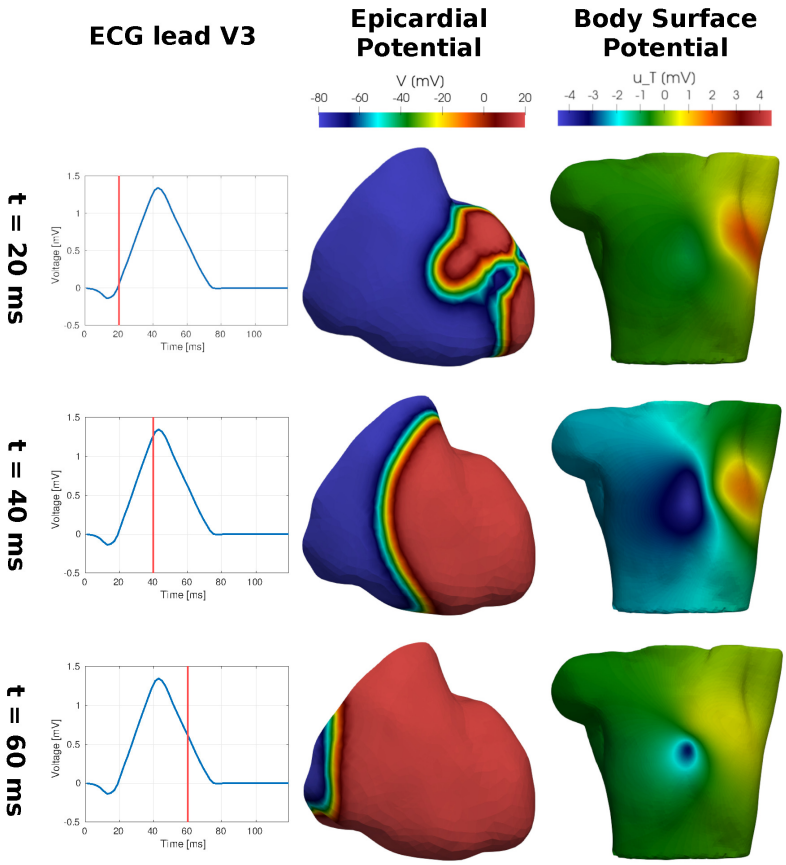

FIG. 7. *Snapshots of one FOM solution to the FPE at $t = 20, 40, 60$ ms. Left: V_3 ECG precordial lead. Center: epicardial extracellular potential field, anterior view. Right: body surface potential field, anterior view.*

of degree $p = 1$. Simulation duration has been set to 160 ms, so that only ventricular depolarization was captured. The computational mesh built on top of the biventricular geometry was made of 221,088 tetrahedral elements, resulting in 49,674 vertices; the mesh constructed on top of the torso geometry, instead, was made of 498,992 tetrahedral cells, which resulted in 94,976 vertices. Concerning temporal discretization, a value of $\Delta t = 0.5$ ms has been chosen. Three snapshots of a FOM solution to the problem are shown in Figure 7.

A crucial step in order to generate a dataset suitable to train DL models is to add sufficient variability. Indeed, on the one side the training datapoints must differ one from the other, so that the model can learn from them the widest possible amount of dynamics and conditions; on the other side, also the testing datapoints should show differences both within each other and with respect to the training ones, so that a proper assessment of model performances can be made. We are aware of the fact that the simplifying modeling assumptions we adopted prevent us from employing our model with success on real data. However, we tried to enrich the data generation process with several sources of randomness, in order to construct a sufficiently variable dataset, able to challenge the capabilities of the ST-RB-DNN model.

TABLE 1
Values of the cardiac electric conductivities employed to numerically approximate heart electrophysiology via the bidomain equations. $\mathcal{U}(a, b)$ denotes a uniform distribution over the interval $[a; b]$.

$\sigma_{\mathbf{l_i}}[S\ cm^{-1}]$	$\sigma_{\mathbf{t_i}}[S\ cm^{-1}]$	$\sigma_{\mathbf{l_e}}[S\ cm^{-1}]$	$\sigma_{\mathbf{t_e}}[S\ cm^{-1}]$
$\mathcal{U}(1.77, 2.92) \cdot 10^{-4}$	$\mathcal{U}(0.66, 1.08) \cdot 10^{-4}$	$\mathcal{U}(1.65, 2.73) \cdot 10^{-3}$	$\mathcal{U}(0.96, 1.32) \cdot 10^{-3}$

Since the FPE consists only of a generalized Laplace equation to be solved in a homogeneous and isotropic volume conductor, no variability has been inserted explicitly at such stage. Regarding heart electrophysiology, instead, variability has been added in two ways. First, all cardiac conductivities (longitudinal and transversal to the fibers' direction, intracellular, and extracellular) have been sampled from a uniform distribution (see Table 1); indeed they have been proved to strongly influence the depolarization process and, in turn, the shape of body surface potentials [4].

Second, different initial activation patterns have been considered; all such patterns are consistent with the findings of Wyndham et al. on epicardial breakthroughs (EBTs) localization in both healthy [56] and LBBB-affected [55] patients. In particular, a thin endocardial and subendocardial layer is activated, for a duration of 5 ms, in 10 possible different regions; this implies that EBTs could occur in 10 possible different locations that mimic the effect of Purkinje fibers (see Figure 5). For healthy patients, all these regions could possibly activate, with the constraint that no more than 5 EBTs should appear. Instead, for patients affected by LBBB (RBBB), only the regions located on the right (left) ventricle can be involved in the stimulation protocol and no more than 3 EBTs should be observed. 50% of the patients in the dataset are healthy, 25% are LBBB-affected, and 25% are RBBB-affected. The position of the point around which the initial stimulation is applied can vary within a sphere of 2 mm radius, obeying a 3D uniform distribution, and the initial activation times are sampled from a uniform distribution, with ranges having been defined according to [56, 55].

Ultimately, the dataset we employed is made of $N_{\boldsymbol{\mu}} = 400$ datapoints. Train and test datasets have been obtained from a 90%−10% splitting of such dataset; the validation dataset has been assembled by picking the 10% of the training datapoints. Furthermore, training and validation datasets have been subject to data augmentation on-the-fly, by superimposing white Gaussian noise proportional to the signals so that an average SNR of 26 dB is achieved. Additionally, starting from the same FOM solutions, two different datasets have been assembled: the first one contains 12-lead ECG signals as body surface signals, while the second one features 158 signals computed from measurements of the body surface potential in 155 different locations (see Figure 6). Both datasets have been suitably modified when the DFT-based model has been considered.

In order to reduce the computational burden of the offline stage of the method, the training and validation datasets have been used also to generate all the RBs. All POD have been performed with respect to the standard ℓ^2-norm and the *randomized SVD* algorithm [23] has been employed:

- The RB in space for the torso potential has been computed by means of a POD applied to the mode-1 unfolding of the corresponding tensor. A tolerance $\epsilon_{POD}^{t,s} = 10^{-3}$ has been chosen and 314 basis functions have been derived.
- The RB in space for the epicardial potential has been computed by means of a POD applied to the mode-1 unfolding of the corresponding tensor. A tolerance $\epsilon_{POD}^{e,s} = 10^{-1}$ has been chosen and 101 basis functions have been derived.

- The dimensionality reduction along the temporal dimension for the epicardial potential has been realized according to the *tailored temporal subspaces via ST-HOSVD* approach proposed in [8], thus computing for each element of the spatial RB a tailored temporal one. The tolerances for the POD $\tilde{\epsilon}_{POD}^{e,t}$ have been set to drop from $5 \cdot 10^{-2}$ to $5 \cdot 10^{-1}$, depending on the singular values associated to the corresponding spatial RB element. Ultimately, the computed space-time RB was made of 766 elements.

A grid search has been conducted both on the time-series-based ST-RB-DNN model and on the DFT-based one, in order to identify optimal values for those hyperparameters having the largest impact on model performance. In particular, all hyperparameters related to the architecture of the trainable encoder have been involved in the grid search and the dataset featuring 12-lead ECG signals was employed. The ℓ^1-norm relative error on the reconstructed epicardial activation maps has been chosen as an evaluation metric. Such maps have been computed by assigning to each vertex of the computational mesh the time instant at which the time derivative of the epicardial extracellular potential was maximal. Results of the grid searches are shown in Table 2 for the time-series-based model and in Table 3 for the DFT-based one. Additionally, Table 4 reports the complexities and the training/testing times for some of the trained models. Concerning other relevant hyperparameters, the learning rate has been set to $\nu = 10^{-3}$ (with a reduction of factor 4 if no improvement of the validation loss could be observed for 20 consecutive epochs); the *Nadam* optimizer (i.e., an improved version of the Adam optimizer incorporating Nesterov momentum) [16] was chosen; ridge regularization has been employed, with parameters $\lambda_r^{FC} = 10^{-7}$ in fully connected layers and $\lambda_r^{Conv} = 10^{-4}$ in convolutional ones. Also, we set $w_{BC} = 10^3$ for the time-series-based model and $w_{BC} = 10^2$ for the DFT-based one; this allowed us to get a loss split of approximately 95%−5% between the error on epicardial potentials and the one on body surface signals. Figures 8 and 9 show the results of the two best models on test datapoint #1 in terms of epicardial activation maps and 12-lead ECG signals, respectively. Figures 10 and 11 show the results of the two best models on test datapoint #3 in terms of epicardial activation maps and 12-lead ECG signals, respectively; these results are relevant, as this datapoint identifies the worst-case scenario for the majority of the trained models. Additionally, Figures 12 and 13 report the results of the two best models on test datapoint #3 in case they are trained on the dataset featuring 158 body surface signals. Table 5 provides a summary of the errors made by the two best models, comparing the performances in case either 12 or 158 body surface signals are provided in input. Figure 14 compares the epicardial activation maps predicted by the best time-series-based and DFT-based models with the ones resulting from the same models trained as pure AEs; in both cases we provided 158 body surface signals in input. For the AEs, no changes in the architecture with respect to the best models have been made, while the additional regularization parameters in (2.21) have been selected upon a grid search process. Their values are $\lambda_t = 10^{-6}$, $\lambda_e = 0$ for the time-series-based AE model and $\lambda_t = 0$, $\lambda_e = 0$ for the DFT-based one. Finally, Table 6 reports the errors made by the best models, trained either in a standard way or as pure AEs.

The training and testing of all models as well as all postprocessing routines have been carried out on a Lenovo ThinkPad T490s mounting Ubuntu 20.04.1 LTS, with 16 GB RAM and an Intel i7-8565U processor with 4 cores at 1.80 GHz. The numerical simulations needed to construct the dataset were instead performed on the *iHEART* cluster (Lenovo SR950 8x24-Core Intel Xeon Platinum 8160, 2.10 GHz and 1.7 TB RAM) at MOX, Dipartimento di Matematica, Politecnico di Milano.

TABLE 2

Average relative errors in ℓ^1-norm on the epicardial activation maps of the test dataset with the time-series-based ST-RB-DNN model. The green cell displays the best model; the red cell displays the worst model. Rows labels are in the form "prelayers - postlayers," where the first entry defines the layers of the preconvolutional fully connected block and the second one the layers of the postconvolutional fully connected block. Column labels are of the form "N_F - K_{dim}," where N_F is the number of convolutional filters and K_{dim} the dimension of the 1D convolutional kernel.

Dense \ **1D Conv**	(5) - 15	(5) - 25	(10) - 15	(10) - 25	(15) - 15	(15) - 25
(32) - (128,64,32,16,8,4)	6.43e-2	6.02e-2	6.52e-2	6.34e-2	6.71e-2	6.06e-2
(64) - (128,64,32,16,8,4)	5.78e-2	6.44e-2	6.95e-2	5.91e-2	8.28e-2	6.15e-2
(32) - (128,64,32,16,8)	5.15e-2	5.44e-2	4.94e-2	5.56e-2	5.40e-2	4.81e-2
(64) - (128,64,32,16,8)	4.51e-2	4.88e-2	5.10e-2	3.96e-2	4.65e-2	5.28e-2
(32) - (128,64,32,16)	5.20e-2	4.09e-2	4.76e-2	4.65e-2	4.32e-2	4.37e-2
(64) - (128,64,32,16)	4.57e-2	4.32e-2	4.23e-2	4.33e-2	4.35e-2	4.43e-2

TABLE 3

Average relative errors in ℓ^1-norm on the epicardial activation maps of the test dataset with the DFT-based ST-RB-DNN model. The green cell displays the best model; the red cell displays the worst model; the yellow cells identify two DFT-based models that will be considered in Table 4. Rows labels are of the form "prelayers - postlayers," the entries corresponding to the number of neurons in the layers preceding and following a Flatten layer. Column labels report the number of DFT coefficients given as input to the model.

Dense \ **DFT modes**	9	17	25	33
(64,64) - (256,256,128,64)	5.64e-2	5.57e-2	5.29e-2	5.46e-2
(128,128) - (256,256,128,64)	5.46e-2	5.58e-2	5.40e-2	4.06e-2
(256,256) - (256,256,128,64)	5.07e-2	5.08e-2	5.20e-2	4.61e-2
(64,64) - (256,256,128)	4.78e-2	4.54e-2	4.14e-2	5.08e-2
(128,128) - (256,256,128)	4.37e-2	4.28e-2	4.26e-2	4.05e-2
(256,256) - (256,256,128)	3.78e-2	3.81e-2	3.67e-2	3.79e-2
(64,64) - (256,256)	4.19e-2	3.91e-2	3.83e-2	4.85e-2
(128,128) - (256,256)	3.90e-2	3.83e-2	4.12e-2	3.90e-2
(256,256) - (256,256)	3.90e-2	3.70e-2	3.77e-2	3.86e-2
(64,64) - (256,512)	3.77e-2	3.79e-2	3.65e-2	5.07e-2
(128,128) - (256,512)	3.60e-2	3.49e-2	3.69e-2	3.75e-2
(256,256) - (256,512)	3.50e-2	3.51e-2	3.55e-2	3.70e-2

TABLE 4

Average activation maps ℓ^1-norm relative error, number of trainable parameters, and training and testing times (in s) for four different ST-RB-DNN models. "Best T-series" and "Best DFT" identify the best time-series-based and DFT-based models, respectively; "DFT Model 1" is a DFT-based model whose complexity is the closest to the one of "Best T-series" among the tested models; "DFT Model 2" is a DFT-based model whose average ℓ^1-norm relative error on activation maps is the closest to the one of "Best T-series" model among the tested models. The architectures of the models can be derived from Tables 2 and 3, following the color-coded notation and the reported activation map errors. T_{train} is the full training time; T_{epoch} is the average time per epoch; $T_{predict}$ is the time required by a single forward pass in the model; T_{test} is the full testing time on a single datapoint, including all postprocessing routines.

	AM error	**# Params**	$\boldsymbol{T_{train}}$	$\boldsymbol{T_{epoch}}$	$\boldsymbol{T_{predict}}$	$\boldsymbol{T_{test}}$
Best T-series	3.96e-2	213′264	603.52 s	11.3871 s	0.0288 s	4.3385 s
Best DFT	3.49e-2	622′590	586.93 s	10.2970 s	0.0329 s	4.6384 s
DFT Model 1	4.78e-2	239′998	446.73 s	9.1180 s	0.0253 s	4.5188 s
DFT Model 2	3.90e-2	358′654	451.92 s	9.2228 s	0.0306 s	4.5453 s

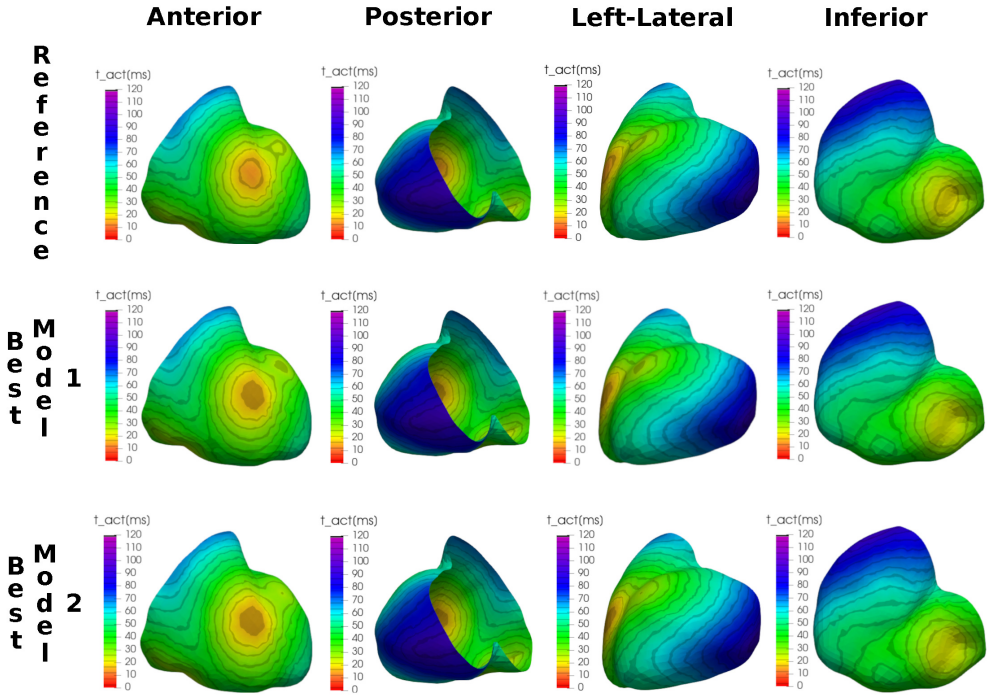

FIG. 8. *Epicardial activation maps for the test datapoint #1 (healthy), computed from three different potential fields: the one reconstructed from the exact ST-ROM approximation (top), the one estimated by the best time-series-based ST-RB-DNN model (center), and the one estimated by the best DFT-based ST-RB-DNN model (bottom). Estimation errors (relative, ℓ^1-norm) are 1.60% and 1.78%, respectively. Both models receive 12-lead ECG signals in input. We show four different views of the activation maps.*

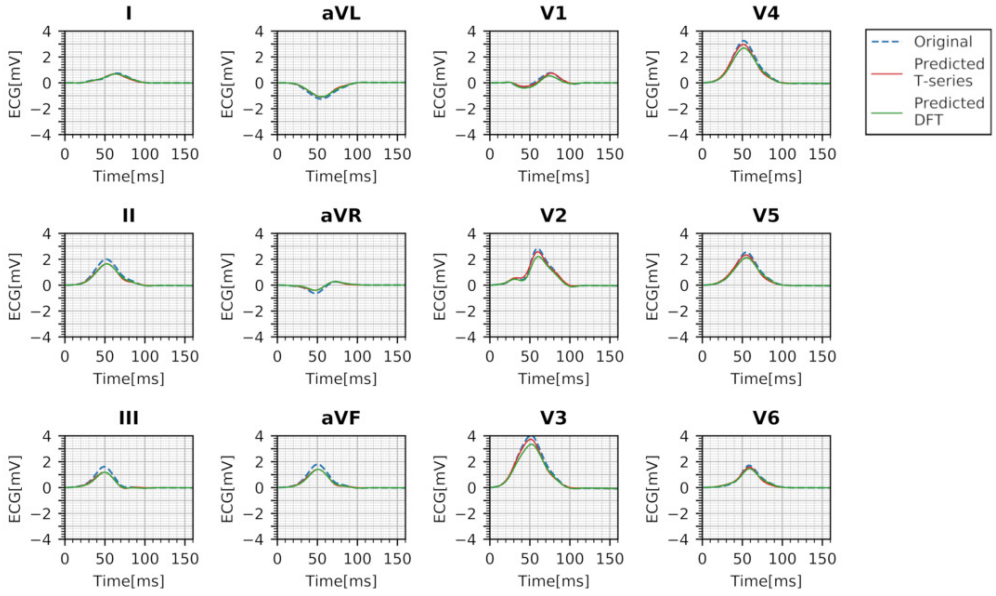

FIG. 9. *12-lead ECG signals for test datapoint #1 (healthy) computed by solving the forward problem with the RB method. Three epicardial potential fields have been employed as Dirichlet boundary datum: the one reconstructed from the exact ST-ROM approximation, the one estimated by the best time-series-based ST-RB-DNN model, and the one estimated by the best DFT-based ST-RB-DNN model. Estimation errors (absolute, ℓ^1-norm, in mV) are $5.63 \cdot 10^{-2}$ and $7.89 \cdot 10^{-2}$, respectively. Both models receive 12-lead ECG signals in input.*

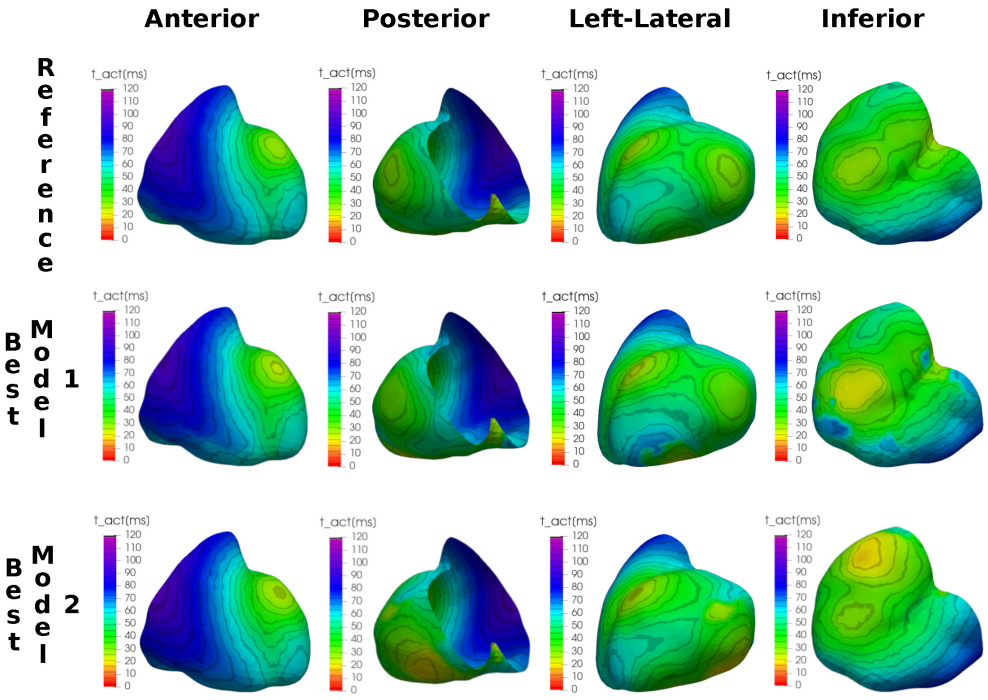

FIG. 10. *Epicardial activation maps for the test datapoint #3 (RBBB), computed from three different potential fields: the one reconstructed from the exact ST-ROM approximation (top), the one estimated by the best time-series-based ST-RB-DNN model (center), and the one estimated by the best DFT-based ST-RB-DNN model (bottom). Estimation errors (relative, ℓ^1-norm) are* 6.36% *and* 9.14%*, respectively. Both models receive* 12*-lead ECG signals in input. We show four different views of the activation maps.*

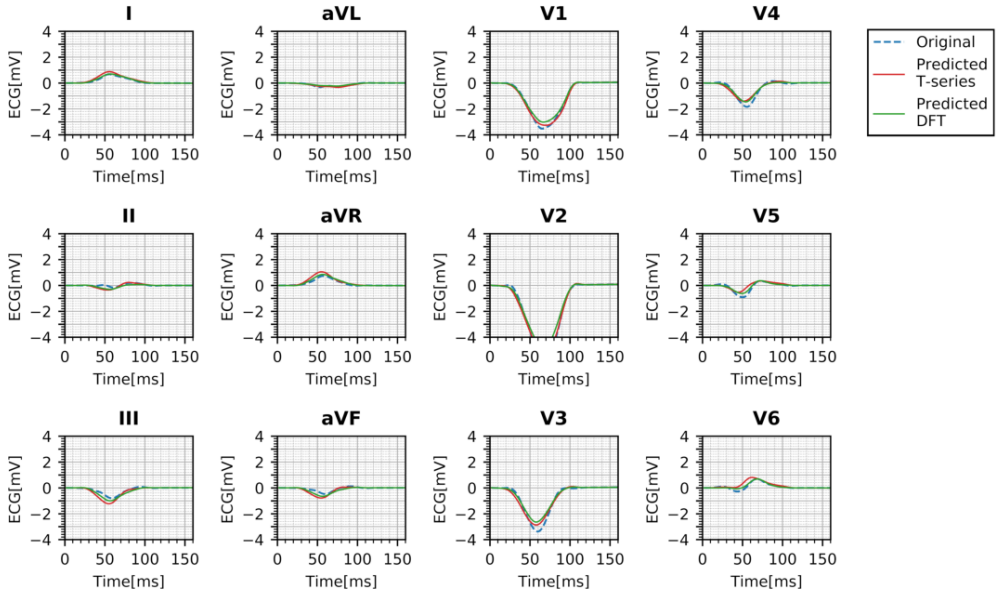

FIG. 11. 12*-lead ECG signals for test datapoint #3 (RBBB) computed by solving the forward problem with the RB method. Three epicardial potential fields have been employed as Dirichlet boundary datum: the one reconstructed from the exact ST-ROM approximation, the one estimated by the best time-series-based ST-RB-DNN, model and the one estimated by the best DFT-based ST-RB-DNN model. Estimation errors (absolute, ℓ^1-norm, in mV) are* $11.51 \cdot 10^{-2}$ *and* $8.23 \cdot 10^{-2}$*, respectively. Both models receive* 12*-lead ECG signals in input.*

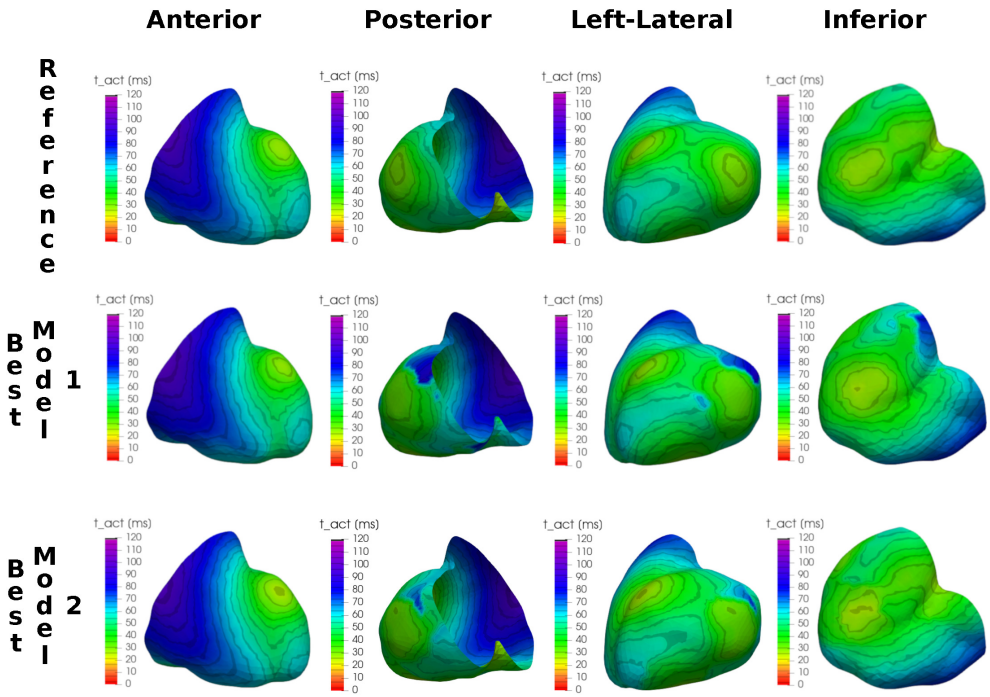

FIG. 12. *Epicardial activation maps for the test datapoint #3 (RBBB), computed from three different potential fields: the one reconstructed from the exact ST-ROM approximation (top), the one estimated by the best time-series-based ST-RB-DNN model (center), and the one estimated by the best DFT-based ST-RB-DNN model (bottom). Estimation errors (relative, ℓ^1-norm) are* 9.26% *and* 4.65%*, respectively. Both models receive* 158 *body surface signals in input. We show four different views of the activation maps.*

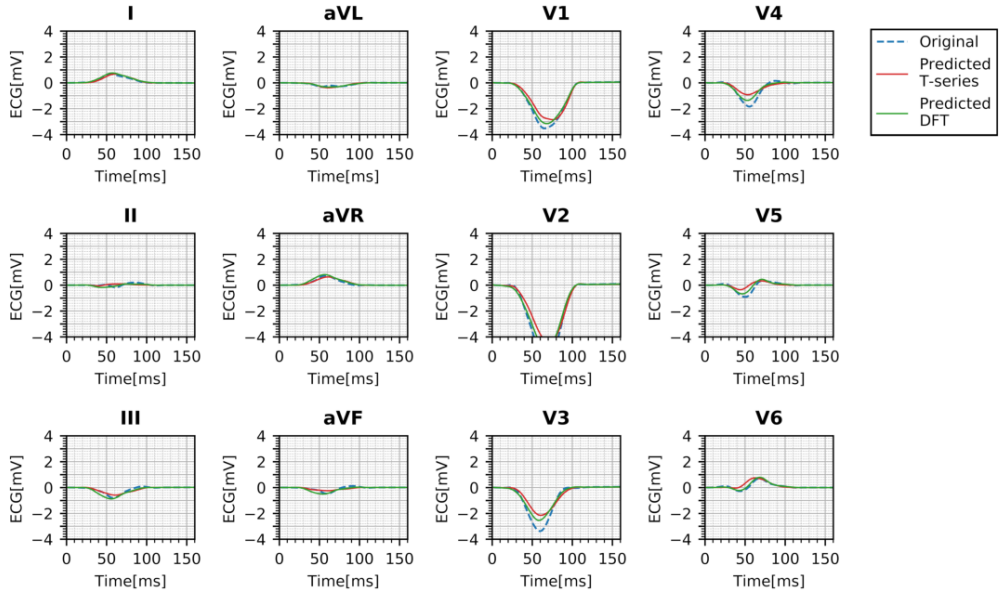

FIG. 13. 12*-lead ECG signals for test datapoint #3 (RBBB) computed by solving the forward problem with the RB method. Three epicardial potential fields have been employed as Dirichlet boundary datum: the one reconstructed from the exact ST-ROM approximation, the one estimated by the best time-series-based ST-RB-DNN model, and the one estimated by the best DFT-based ST-RB-DNN model. Estimation errors (absolute, ℓ^1-norm, in mV) are* $11.71 \cdot 10^{-2}$ *and* $8.19 \cdot 10^{-2}$*, respectively. Both models receive* 158 *body surface signals in input.*

TABLE 5

Reconstruction errors on the epicardial activation maps (left) and on 12-lead ECG signals (right) made by the best model, depending on the nature and on the dimensionality of the input. We measure activation maps' errors and ECG signals' errors with relative ℓ^1-norm and absolute (in mV) ℓ^1-norm, respectively.

Size \ Type	TS	DFT
12	3.96e-2	3.49e-2
158	3.98e-2	3.29e-2

Size \ Type	TS	DFT
12	6.52e-2 mV	6.82e-2 mV
158	7.55e-2 mV	6.82e-2 mV

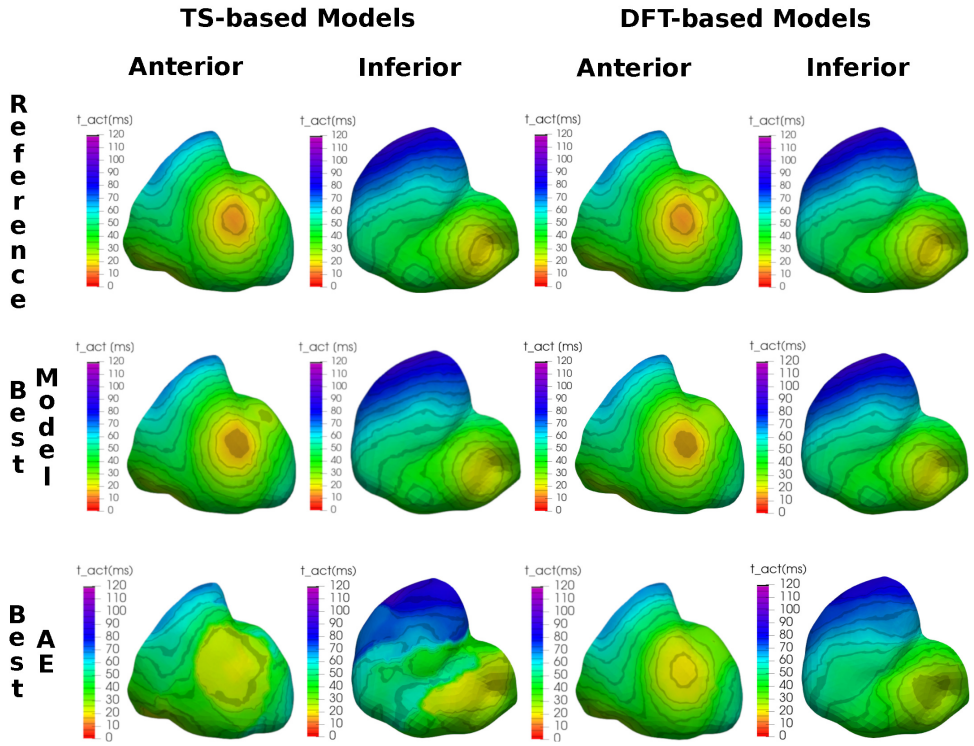

FIG. 14. *Epicardial activation maps for the test datapoint #1 (healthy), computed from five different potential fields:* (i) *the one reconstructed from the exact ST-ROM approximation (top),* (ii) *the one estimated by the best time-series-based ST-RB-DNN model (center-left),* (iii) *the one estimated by the best DFT-based ST-RB-DNN model (center-right),* (iv) *the one estimated by the best time-series-based ST-RB-DNN AE model (bottom-left), and* (v) *the one estimated by the best DFT-based ST-RB-DNN AE model (bottom-right). Estimation errors (relative, ℓ^1-norm) are* 1.60% *and* 1.78% *for the non-AE models and* 11.88% *and* 5.82% *for the AE models. All models receive* 158 *body surface signals in input. We show two different views of the activation maps.*

TABLE 6

Reconstruction errors on the epicardial activation maps (left) and on 12-lead ECG signals (right) made by the best model, depending on the nature of the input (TS or DFT) and on the type of training (not-AE or AE). We measure activation maps' errors and ECG signals' errors with relative ℓ^1-norm and absolute (in mV) ℓ^1-norm, respectively.

Train \ Type	TS	DFT
not-AE	3.98e-2	3.29e-2
AE	14.49e-2	8.81e-2

Train \ Type	TS	DFT
not-AE	7.55e-2 mV	6.82e-2 mV
AE	6.33e-2 mV	4.15e-2 mV

4. Discussion. Our study focused on the development of a DL model (called ST-RB-DNN) that allows to estimate physically consistent solutions to the inverse problem of electrophysiology both leveraging data availability and exploiting the awareness of the physical laws governing the phenomenon of interest, expressed by means of suitable PDEs.

4.1. Main findings. In light of the obtained numerical results, the main findings are listed in the following.

The model provides reliable estimations of epicardial activation maps. The time-series-based ST-RB-DNN model trained with 12-lead ECG signals in input presents good performance in terms of accuracy. Table 2 shows the results of the grid search process on architecture-related hyperparameters (in terms of the average ℓ^1-norm relative error on the test set); the table is organized so that model complexity increases both rowwise (top to bottom) and columnwise (left to right). Remarkably, none of the considered models features an error higher than 10% and the best one, with an error of 3.96%, presents an intermediate level of complexity (213,264 trainable parameters; see Table 4), proving that no overfitting did occur. Again referring to Table 4, we can notice that the training time equals ≈604 s (11.4 s per epoch on average); testing on one fresh data sample, instead, takes 4.34 s, of which just 0.029 s are due to the forward pass in the model, while the others are related to the postprocessing routines that allows to reconstruct the FOM solution and to compute the epicardial activation map. Activation maps are qualitatively well reconstructed by the "best model" with a good identification of both EBT and latest epicardial activation (LEA) areas (see Figures 8 and 10, second row). The level of accuracy we attain allows then to capture a vast part of the variability of the phenomenon at hand, comprising both physiological and pathological (in our case, limited to LBBB and RBBB conduction defects) conditions. Additionally, body surface signals are well approximated in the autoencoding portion of the output (see Figures 9 and 11, red solid line), with an average ℓ^1-norm absolute error on the test dataset of just $6.52 \cdot 10^{-2}$ mV; the presence of an embedded RB solver of the FPE as deterministic decoder happens to play a key role on this aspect.

Signals preprocessing via DFT improves model performances. If the ST-RB-DNN model is adapted so that it can take as input the lowest-frequency coefficients arising from the application of a DFT to 12-lead ECG signals, its performance improves. Table 3 reports the results of the grid search process on architecture-related hyperparameters and on the number of DFT coefficients; the table is organized so that model complexity increases both rowwise and columnwise. Also in this case, all average ℓ^1-norm relative errors are below the threshold of 10%; moreover, almost all models feature an error that is lower than 5% and that equals 3.49% for the best model (−0.47% with respect to the best time-series-based model). In this case, the time required by testing on a fresh data sample equals 4.64 s, of which only 0.033 s are related to the forward pass in the model (see Table 4). About computational complexity, on the one side lower errors are achieved if a limited number of DFT coefficients is given in input. Indeed providing information on high-frequency modes happens to be not only useless, but even counterproductive, as ultimately a spatio-temporal reduced representation of the epicardial potential is estimated. On the other side, the best model has 622,590 trainable parameters and, fixing the input dimensionality, it is the second most complex among the considered ones. By proceeding rowwise along each column of Table 3, it can be noticed that errors seem to reach a plateau, so that no significant improvements should be expected by further increasing the number of hyperparameters; no additional tests have been made on this point. Noticeably,

despite featuring a much higher complexity with respect to the best time-series-based model, the best DFT-based one could be trained in a slightly lower amount of time (see Table 4); this is due to the fact that the trainable encoder of such a model is no longer a 1D CNN but a simple MLP, hence the backpropagation algorithm can be performed much faster. Also, from Table 4 it can be noticed that a DFT-based model as complex as the best time-series-based one features a much higher error (4.78% versus 3.96%), while a DFT-based model whose performance is comparable with that of the best time-series-based model exhibits a higher complexity (358,654 parameters versus 213,264 parameters), but it could be trained in a much lower amount of time ($\approx$452 s versus $\approx$604 s). As expected, the epicardial activation maps estimated by the best DFT-based model are qualitatively close to the target ones (see Figures 8 and 10, third row), showing a good identification of both EBT and LEA locations. Actually, in the worst-case scenario presented in Figure 10, the DFT-based model happens to estimate a nonphysical EBT in the infero-lateral basal portion of the left ventricle; such misidentifications have been encountered only in a minority of the 40 testing datapoints (for both models) and they were always associated to low-occurrence EBTs. For this reason, we expect the problem to be avoided by either just increasing the dimensionality of the training dataset or upsampling the critical EBTs at the data generation stage. Finally, body surface signals have been well reconstructed at all leads in both reported test datapoints (see Figures 9 and 11, green solid line), with an average ℓ^1-norm absolute error on the test dataset equaling $6.82 \cdot 10^{-2}$ mV ($+0.30 \cdot 10^{-2}$ mV with respect to the best time-series-based model).

Considering more body surface signals does not severely impact on model performance. The best model architectures, trained either with 158 body surface signals or with 12-lead ECG ones, exhibited comparable performance. In particular, the time-series-based model featured an average ℓ^1-norm relative error of 3.98% (+0.02%), while that drops to 3.29% for the DFT-based one (-0.20%). The reduced dimensionality and heterogeneity of the training dataset is thought to exert a big impact on this aspect, as 12 body surface signals may be enough to encode the variability of the considered data. Thus, the effective usefulness of providing more body surface signals—common in ECGI clinical applications—should be assessed when dealing with bigger datasets. This consideration is enhanced by looking at Figure 12, which reports the epicardial activation maps estimated by the best models taking 158 signals in input: indeed it can be noticed that the DFT-based model no longer estimates the nonphysical EBT on the infero-lateral basal portion of the left ventricle, which was instead detected by the corresponding model taking 12-lead ECG signals in input (see Figure 10).

Training the model as a pure AE worsens its performance. When the best model architectures (with 158 body surface signals in input) are trained as pure AEs, the prediction accuracy on epicardial activations maps worsens (see Table 6). In particular, for the time-series-based model the average ℓ^1-norm relative error increases from 3.98% to 14.49%, while for the DFT-based one it rises from 3.29% to 8.81%. We can also qualitatively visualize the decay of the model performance in Figure 14. Indeed the epicardial activation maps reconstructed from the solutions predicted by the AE models appear less accurate than the ones resulting from the non-AE models. This is particularly evident for the time-series-based model. In the DFT-based one, instead, all epicardial EBT locations are well identified, but the EBT timings prediction is less accurate. These results show that signal preprocessing via DFT dramatically improves model performance when we train it as a pure AE. However we remark that no grid search on architectural hyperparameters has been conducted on such models;

therefore, the much higher complexity of the DFT-based model with respect to the time-series-based one (712,190 versus 222,190 trainable parameters) may also explain the larger performance gap. We also notice that the average ℓ^1-norm absolute errors on 12-lead ECG signals diminish if the models are trained as pure AEs, dropping from $7.55 \cdot 10^{-2}$ mV to $6.33 \cdot 10^{-2}$ mV for the time-series-based model and from $6.82 \cdot 10^{-2}$ mV to $4.15 \cdot 10^{-2}$ mV for the DFT-based one. This result is expected since the error on body surface signals becomes the only term in the loss (aside of regularizations) if the models are trained as pure AEs. Finally, the additional regularization terms (see (2.21)) do not impact model performance. In particular the L^2-norm regularization on epicardial potentials does not help in reducing errors (indeed $\lambda_e = 0$ for both models), while the penalization of the H^1-norm of the torso potential allows us to slightly ameliorate the predictions only in the time-series-based case ($\lambda_t = 10^{-6}$).

Exploiting physical awareness allows constructing low-complexity and fast-training models in a small data regime. As reported in Table 4, all the implemented ST-RB-DNN models, despite the small amount of synthetic data at our disposal, managed to provide good estimations of the epicardial potential field during ventricular depolarization and could be trained on a simple laptop in ≈ 15 min. The two elements that allowed us to get physical awareness play a key role in this:

- The projection of the epicardial potential onto a space-time reduced subspace, generated from the same physical FOM solutions that have been employed to construct the training dataset, dramatically reduces the number of coefficients to be estimated by the trainable encoder, other than allowing us to achieve independence from the space-time mesh refinement and to constrain the solution to a lower-dimensional and physically consistent manifold. In this way, the training process of the model, via a backpropagation algorithm, becomes feasible, while it would have been extremely difficult, as well as very expensive, if FOM approximations of the epicardial potential field had had to be predicted (see [31] for additional details). In addition, the reduced dimensionality of the dataset allowed us to train the models on a simple laptop. If much more data were at our disposal (as would be necessary for effective clinical validation), we would need more computational resources and a more optimized training pipeline. However, the trainability of the models is not expected to change in any way.
- The insertion of an embedded RB solver for the FPE as a deterministic decoder and its employment as a physically informed regularization agent in the expression of the loss functional (see (2.19)) allowed us to design a physics-based NN, similar to the ones presented in [15]. As stated in [40], building ML models able to exploit knowledge of the most relevant physical laws governing the phenomenon of interest proved to be effective in the *small data regime*, i.e., when the amount of data at disposal is limited due to the high cost/complexity of the acquisition procedures. In fact, just starting from a small dataset generated *in silico*, we managed to train low-complexity models, showing a remarkable representation power.

4.2. Limitations. The current work represents a first methodological attempt of tackling the inverse problem of electrocardiography by combining knowledge of the main physical laws governing the problem at hand, expressed via suitable PDEs/ODEs, with DL techniques for the exploitation of data availability. Configuring as a preliminary analysis, some simplifying assumptions have been made and they should be addressed in order to bridge the gap with clinical applications. The most relevant limitations are the following.

Simplified dataset. The ST-RB-DNN model training necessitates many epicardial activation maps. Unfortunately, acquiring such maps from clinical measurements requires invasive procedures. To circumvent this shortcoming, we trained and tested our models with *in silico* data, obtained by numerically approximating the solutions to heart electrophysiology and to the FPE. Due to the high computational costs required by detailed FOM simulations, we made some simplifying assumptions as the employment of not highly refined computational meshes and the approximation of the torso as a homogeneous and isotropic volume conductor. Furthermore, we imposed a one-way coupling between the two aforementioned problems via the *isolated heart assumption*, discarding the continuity of the currents at the epicardial surface. Despite appearing relevant from the physical point of view, this last assumption is not expected to seriously impact on our results. Indeed, exposing the heart to insulating air has been proven, both numerically [4] and experimentally [22], to determine an increase in the epicardial potential magnitudes, but without affecting in most cases the activation pattern.

Reference geometry. All the training and testing datapoints have been generated on the same reference geometry, preventing us from taking into account geometrical variability. The position and the rotation of the heart inside the torso can vary considerably, both from patient to patient and, for the same patient, from heartbeat to heartbeat, having a nonnegligible impact on body surface signals. These effects could be taken into account by using methods that allow tracking the position of the heart from body surface potentials, as the one presented in [12].

White noise. Another key aspect that differentiates clinically measured body surface potentials from numerically approximated ones is the presence of noise. We have taken this element into consideration by superimposing white Gaussian noise to the simulated signals; the noise has been chosen as proportional to the signals themselves at each discrete time instant and an average SNR of $\approx$26 dB has been obtained. However, real body surface potentials feature the presence of different sources of noise, linked to various aspects (muscular contraction/relaxation, breathing, measurement instruments and procedures, etc.) and whose distribution is not necessarily white and Gaussian (see [44, 45]). Thus, an interesting development toward clinical application would be to superimpose to the numerically generated signals more realistic noise components, as done, for instance, in [17].

Limited design research. In terms of NN design, only a few architectures have been tested for the encoder. Actually, during the first stages of development, we also considered other architectures such as long-short term memory networks [25], residual networks [29], or convolutional recurrent neural networks [58]; however, as no significant improvements with respect to CNN and MLP were observed, no further investigations have been made in this direction. Different choices in terms of architectural design and optimization algorithms may allow us to ease the constraints on the input/output datasets and eventually lead to better model performance.

5. Conclusion. The present study concerned the development of a physics-informed deep learning model, named ST-RB-DNN, able to provide physically consistent and data-driven solutions to the inverse problem of electrocardiography. In the numerical test cases that have been conducted, all variants of the model (distinguished based on the dimensionality and nature of the provided body surface signals) proved to be accurate, up to an average error of $\approx$3.5% in relative ℓ^1-norm, considering the epicardial activation maps of the test dataset. Additionally, despite working under several simplifying assumptions and employing a small amount of *in silico* data for both training and testing, our models proved anyway to be highly efficient; their

trainings could all be performed in ≈15 min on a simple laptop and a single forward pass only required ≈0.03 s. To conclude, the obtained results indicate that our method has potential to be used in clinical practice, notwithstanding some improvements, mostly concerning the data generation stage.

Appendix A. Abbreviations and notation.

A.1. List of abbreviations.

AE: Autoencoder
AP: Aliev–Panfilov
BSPM: Body surface potential maps
CNN: Convolutional neural network
CT: Computed tomography
CRNN: Convolutional recurrent neural network
CVA: Conditional variational autoencoder
DFT: Discrete Fourier transform
DL: Deep learning
DNN: Deep neural network
DOF: Degree of freedom
EBT: Epicardial breakthrough
ECG: Electrocardiography
ECGI: Electrocardiographic Imaging
ELU: Exponential linear unit
FE: Finite element
FFT: Fast Fourier transform
FHN: FitzHugh–Nagumo
FOM: Full order model
FPE: Forward problem of electrocardiography
HOSVD: High order singular value decomposition
LBBB: Left bundle branch block
LEA: Latest epicardial activation
LSTM: Long-short term memory
MAE: Mean absolute error
ML: Machine learning
MLP: Multiple layer perceptron
MSE: Mean squared error
NN: Neural network
ODE: Ordinary differential equation
PDE: Partial differential equation
PINN: Physically informed neural network
POD: Proper orthogonal decomposition
RB: Reduced basis
RBBB: Right bundle branch block
ReLU: Rectified linear unit
ResNet: Residual network
ROM: Reduced order model
SELU: Scaled exponential linear unit
SNR: Signal-to-noise ratio
ST: Space-time (or spatio-temporal)
SVD: Singular value decomposition
TS: Time series

A.2. Summary of notation.

1. **Functional spaces** (assuming $\Omega \subset \mathbb{R}^n$ is a Lipschitz domain)
   - $L^2(\Omega) = \left\{u : \Omega \to \mathbb{R} \;\; s.t. \;\; \int_\Omega |u|^2 \, dx < \infty\right\}$
   - $H^1(\Omega) = \{u \in L^2(\Omega) \;\; s.t. \; \frac{\partial u}{\partial x_i} \in L^2(\Omega) \; \forall i \in \{1, \dots, n\}\}$
   - $H^{\frac{1}{2}}(\partial\Omega) = \left\{u \in L^2(\partial\Omega) \;\; s.t \; \exists \tilde{u} \in H^1(\Omega) : Tr(\tilde{u}) = u\right\}$ where $Tr : H^1(\Omega) \to L^2(\partial\Omega)$ the trace operator.
2. **Norms** (assuming $\Omega \subset \mathbb{R}^n$ is a Lipschitz domain)
   - $||u||_{L^2(\Omega)} = \left(\int_\Omega |u|^2 \, dx\right)^{\frac{1}{2}}$
   - $|u|_{H^1(\Omega)} = ||\nabla u||_{[L^2(\Omega)]^n} = \left(\int_\Omega |\nabla u|^2 \, dx\right)^{\frac{1}{2}}$
   - $||u||_{H^1(\Omega)} = \left(||u||^2_{L^2(\Omega)} + |u|^2_{H^1(\Omega)}\right)^{\frac{1}{2}}$
3. **Geometry and time**
   - Ω_H: heart domain
   - Ω_T: torso domain
   - Γ_B: body surface boundary
   - Γ_H: epicardial surface boundary
   - t_0: initial time
   - T: final time
4. **Physical models of cardiac electrophysiology and FPE**
   - u_i: intracellular potential
   - u_e: extracellular potential
   - $v =: u_i - u_e$: transmembrane potential
   - w: gating variable of the AP ionic model
   - u_T: torso potential
   - $\boldsymbol{D}_i$: intracellular conductivity tensor
   - $\boldsymbol{D}_e$: extracellular conductivity tensor
   - $\boldsymbol{D}_T$: torso conductivity tensor
   - A_m: cell membrane area per unit volume
   - C_m: membrane capacitance per unit area
   - I_{app}: externally applied current
   - I_{ion}: ionic current
   - $\sigma_{l_{i,e}}$: longitudinal electrical conductivity in the intracellular/extracellular domain
   - $\sigma_{t_{i,e}}$: transversal electrical conductivity in the intracellular/extracellular domain
5. **FOMs for cardiac electrophysiology and FPE**
   - N_h^e: dimension of the FE space for the extracellular and transmembrane potentials
   - N_h: dimension of the FE space for the torso potential
   - N_t: number of time steps
   - $\boldsymbol{u}_{e_h}$: discretized extracellular potential
   - $\boldsymbol{v}_h$: discretized transmembrane potential
   - $\boldsymbol{w}_h$: discretized gating variable
   - $\boldsymbol{u}_{t_h}$ discretized torso potential
   - $\boldsymbol{u}_{e_h}^{\Gamma_H}$: discretized trace of the extracellular potential on the epicardial surface Γ_H
   - $\boldsymbol{A}_{in}$: stiffness matrix for the intracellular domain
   - $\boldsymbol{A}_{ex}$: stiffness matrix for the extracellular domain
   - $\boldsymbol{M}$: mass matrix for the biventricular domain

- $\boldsymbol{A}$: stiffness matrix for the torso domain
- $\boldsymbol{A}_{\Gamma_H}$: stiffness matrix for the torso domain, restricted to the epicardial surface

6. **Model order reduction of the FPE**
   - $N_{\boldsymbol{\mu}}$: number of snapshots (i.e., parameter values) for the offline phase
   - $\mathcal{X}$: tensor of the snapshots for the torso potential
   - $\mathcal{X}^e$: tensor of the snapshots for the epicardial potential
   - $\boldsymbol{V}_s$: matrix encoding the reduced basis in space for the torso potential
   - $\boldsymbol{V}_s^e$: matrix encoding the reduced basis in space for the epicardial potential
   - n_h: number of spatial basis functions for the torso potential
   - n_h^e: number of spatial basis functions for the epicardial potential
   - $\tilde{\boldsymbol{u}}_{t_h}$: space-reduced torso potential
   - $\tilde{\boldsymbol{u}}_{e_h}^{\Gamma_H}$: space-reduced trace of the extracellular potential on the epicardial surface Γ_H
   - $\tilde{\boldsymbol{A}}_{\boldsymbol{\mu}}$: reduced stiffness matrix for the torso domain
   - $\tilde{\boldsymbol{A}}^e$: reduced stiffness matrix for the torso domain, restricted to the epicardial surface
   - $\boldsymbol{V}_{t_i}^e$: matrix encoding the temporal basis functions for the epicardial potential, tailored to the ith spatial basis function
   - n_t^i: number of temporal basis functions for the epicardial potential, tailored to the ith spatial basis function
   - n_{st}: number of spatio-temporal basis functions for the epicardial potential
   - $\hat{\boldsymbol{u}}_{e_h}^{\Gamma_H}$: space-time-reduced trace of the extracellular potential on the epicardial surface Γ_H
   - $\mathcal{F}$: mapping from the space and time reduced basis indices to the space-time reduced basis index for the epicardial potential
7. **ST-RB-DNN model**
   - $\boldsymbol{\Theta}$: vector of the model hyperparameters
   - b_s: batch size employed during model training
   - n_{sig}: number of input/output body surface signals
   - $\bar{\boldsymbol{u}}_{e_h}^{a,\boldsymbol{\mu}}$: estimated latent representation of epicardial potential for the parameter value $\boldsymbol{\mu}$, prior to the application of the affine transformation
   - $\hat{\boldsymbol{u}}_{e_h}^{a,\boldsymbol{\mu}}$: estimated space-time-reduced epicardial potential for the parameter value $\boldsymbol{\mu}$
   - $\hat{\boldsymbol{u}}_{e_h}^{r,\boldsymbol{\mu}}$: target space-time-reduced epicardial potential for the parameter value $\boldsymbol{\mu}$
   - $\tilde{\boldsymbol{u}}_{e_h}^{a,\boldsymbol{\mu}}$: estimated space-reduced epicardial potential for the parameter value $\boldsymbol{\mu}$
   - $\tilde{\boldsymbol{u}}_{t_h}^{a,\boldsymbol{\mu}}$: estimated space-reduced torso potential for the parameter value $\boldsymbol{\mu}$
   - $\boldsymbol{S}_{\#i}^{r,\boldsymbol{\mu}}$: target ith body surface signal for the parameter value $\boldsymbol{\mu}$
   - $\boldsymbol{S}_{\#i}^{a,\boldsymbol{\mu}}$: estimated ith body surface signal for the parameter value $\boldsymbol{\mu}$
   - $\hat{\boldsymbol{u}}_{e_h}^{shift}$: additive term in the affine transformation of space-time-reduced epicardial potentials
   - $\hat{\boldsymbol{u}}_{e_h}^{scale}$: scaling term in the affine transformation of space-time-reduced epicardial potentials
   - $\mathcal{L}(\boldsymbol{\Theta})$: loss function
   - $\mathcal{L}_{BC}(\boldsymbol{\Theta})$: loss function contribution due to the epicardial potentials
   - $\mathcal{L}_{sig}(\boldsymbol{\Theta})$: loss function contribution due to the body surface signals

- $\mathcal{L}_{reg}(\boldsymbol{\Theta})$: loss function contribution due to the regularization
- $\mathcal{L}_{reg}^{AE}(\boldsymbol{\Theta})$: loss function contribution due to the regularization in ST-RB-DNN AEs
- w_{BC}: weight of the error on space-time-reduced epicardial potentials in the loss
- λ_r: weight of the ridge regularization term in the loss
- λ_t: weight of the torso potential regularization term in the loss for ST-RB-DNN AEs
- λ_r: weight of the epicardial potential regularization term in the loss for ST-RB-DNN AEs
- $\sigma_j^{e,s}$: jth singular value arising from the spatial POD applied to the epicardial potential tensor $\mathcal{X}^e$
- $\sigma_k^{e,t,j}$: kth singular value arising from the temporal POD applied to the projection of the epicardial potential tensor $\mathcal{X}^e$ onto the 1D subspace spanned by the jth spatial reduced basis function

REFERENCES

[1] R. R. Aliev and A. V. Panfilov, *A simple two-variable model of cardiac excitation*, Chaos Solitons Fractals, 7 (1996), pp. 293–301.

[2] T. Bacoyannis, J. Krebs, N. Cedilnik, H. Cochet, and M. Sermesant, *Deep learning formulation of ECGI for data-driven integration of spatiotemporal correlations and imaging information*, in International Conference on Functional Imaging and Modeling of the Heart, Springer, New York, 2019, pp. 20–28.

[3] L. R. Bear, L. K. Cheng, I. J. LeGrice, G. B. Sands, N. A. Lever, D. J. Paterson, and B. H. Smaill, *Forward problem of electrocardiography: Is it solved?*, Circulation Arrhythmia Electrophysiology, 8 (2015), pp. 677–684.

[4] M. Boulakia, S. Cazeau, M. A. Fernández, J.-F. Gerbeau, and N. Zemzemi, *Mathematical modeling of electrocardiograms: a numerical study*, Ann. Biomedical Engineering, 38 (2010), pp. 1071–1097.

[5] K. Bujnarowski, P. Bonizzi, M. Cluitmans, R. Peeters, and J. Karel, *CT-scan free neural network-based reconstruction of heart surface potentials from ECG recordings*, Comput. Cardiology, 47 (2020).

[6] S. Butterworth et al., *On the theory of filter amplifiers*, Wireless Engineer, 7 (1930), pp. 536–541.

[7] S. Cai, Z. Wang, L. Lu, T. A. Zaki, and G. E. Karniadakis, *DeepM&Mnet: Inferring the Electroconvection Multiphysics Fields Based on Operator Approximation by Neural Networks*, preprint, https://arxiv.org/abs/2009.12935, 2020.

[8] Y. Choi and K. Carlberg, *Space-time least-squares Petrov–Galerkin projection for nonlinear model reduction*, SIAM J. Sci. Comput., 41 (2019), pp. A26–A58.

[9] M. Cluitmans, D. H. Brooks, R. MacLeod, O. Dössel, M. S. Guillem, P. M. van Dam, J. Svehlikova, B. He, J. Sapp, L. Wang, et al., *Validation and opportunities of electrocardiographic imaging: From technical achievements to clinical applications*, Frontiers in Physiology, 9 (2018), 1305.

[10] M. J. Cluitmans, P. Bonizzi, J. M. Karel, M. Das, B. L. Kietselaer, M. M. de Jong, F. W. Prinzen, R. L. Peeters, R. L. Westra, and P. G. Volders, *In vivo validation of electrocardiographic imaging*, JACC Clinical Electrophysiology, 3 (2017), pp. 232–242.

[11] M. J. M. Cluitmans, R. Peeters, R. Westra, and P. Volders, *Noninvasive reconstruction of cardiac electrical activity: Update on current methods, applications and challenges*, Netherlands Heart J., 23 (2015), pp. 301–311.

[12] J. Coll-Font and D. H. Brooks, *Tracking the position of the heart from body surface potential maps and electrograms*, Frontiers Physiology, 9 (2018), 1727.

[13] P. Colli Franzone and L. F. Pavarino, *A parallel solver for reaction–diffusion systems in computational electrocardiology*, Math. Models Methods Appl. Sci., 14 (2004), pp. 883–911.

[14] J. W. Cooley and J. W. Tukey, *An algorithm for the machine calculation of complex Fourier series*, Math. Comp., 19 (1965), pp. 297–301.

[15] N. Dal Santo, S. Deparis, and L. Pegolotti, *Data driven approximation of parametrized PDEs by Reduced Basis and Neural Networks*, J. Comput. Phys., 416 (2020), 109550.

[16] T. Dozat, *Incorporating Nesterov Momentum into Adam*, http://cs229.stanford.edu/proj2015/054_report.pdf, 2016.
[17] O. El B'charri, R. Latif, K. Elmansouri, A. Abenaou, and W. Jenkal, *ECG signal performance de-noising assessment based on Threshold tuning of dual-tree wavelet transform*, Biomedical Engineering Online, 16 (2017), 26.
[18] A. Ferrer, R. Sebastián, D. Sanchez-Quintana, J. F. Rodriguez, E. J. Godoy, L. Martinez, and J. Saiz, *Detailed anatomical and electrophysiological models of human atria and torso for the simulation of atrial activation*, PloS ONE, 10 (2015), e0141573.
[19] P. C. Franzone, L. F. Pavarino, and S. Scacchi, *Mathematical Cardiac Electrophysiology*, MS&A 13, Springer, New York, 2014.
[20] S. Giffard-Roisin, T. Jackson, L. Fovargue, J. Lee, H. Delingette, R. Razavi, N. Ayache, and M. Sermesant, *Noninvasive personalization of a cardiac electrophysiology model from body surface potential mapping*, IEEE Trans. Biomedical Engineering, 64 (2016), pp. 2206–2218.
[21] E. Grandelius and K. H. Karlsen, *The Cardiac Bidomain Model and Homogenization*, preprint, https://arxiv.org/abs/1811.07524, 2018.
[22] L. Green, B. Taccardi, P. Ershler, and R. Lux, *Epicardial potential mapping. Effects of conducting media on isopotential and isochrone distributions.*, Circulation, 84 (1991), pp. 2513–2521.
[23] N. Halko, P.-G. Martinsson, and J. A. Tropp, *Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions*, SIAM Rev., 53 (2011), pp. 217–288.
[24] J. S. Hesthaven and S. Ubbiali, *Non-intrusive reduced order modeling of nonlinear problems using neural networks*, J. Comput. Phys., 363 (2018), pp. 55–78.
[25] S. Hochreiter and J. Schmidhuber, *Long short-term memory*, Neural Comput., 9 (1997), pp. 1735–1780.
[26] C. Hoogendoorn, N. Duchateau, D. Sanchez-Quintana, T. Whitmarsh, F. M. Sukno, M. De Craene, K. Lekadir, and A. F. Frangi, *A high-resolution atlas and statistical model of the human heart from multislice CT*, IEEE Trans. Medical Imaging, 32 (2012), pp. 28–44.
[27] A. Intini, R. N. Goldstein, P. Jia, C. Ramanathan, K. Ryu, B. Giannattasio, R. Gilkeson, B. S. Stambler, P. Brugada, W. G. Stevenson, et al., *Electrocardiographic imaging (ECGI), a novel diagnostic modality used for mapping of focal left ventricular tachycardia in a young athlete*, Heart Rhythm, 2 (2005), pp. 1250–1252.
[28] E. M. Izhikevich and R. FitzHugh, *FitzHugh-Nagumo model*, Scholarpedia, 1 (2006), 1349.
[29] M. Kachuee, S. Fazeli, and M. Sarrafzadeh, *ECG heartbeat classification: A deep transferable representation*, in Proceedings of the International Conference on Healthcare Informatics, IEEE, 2018, pp. 443–444.
[30] A. Karoui, L. Bear, P. Migerditichan, and N. Zemzemi, *Evaluation of fifteen algorithms for the resolution of the electrocardiography imaging inverse problem using ex-vivo and in-silico data*, Frontiers Physiology, 9 (2018), 1708.
[31] G. Kutyniok, P. Petersen, M. Raslan, and R. Schneider, *A Theoretical Analysis of Deep Neural Networks and Parametric PDEs*, preprint, https://arxiv.org/abs/1904.00377, 2019.
[32] L. Lu, P. Jin, and G. E. Karniadakis, *Deeponet: Learning Nonlinear Operators for Identifying Differential Equations Based on the Universal Approximation Theorem of Operators*, preprint, https://arxiv.org/abs/1910.03193, 2019.
[33] L. Lu, Y. Shin, Y. Su, and G. E. Karniadakis, *Dying Relu and Initialization: Theory and Numerical Examples*, preprint, https://arxiv.org/abs/1903.06733, 2019.
[34] B. Messinger-Rapport and Y. Rudy, *Noninvasive recovery of epicardial potentials in a realistic heart-torso geometry. Normal sinus rhythm*, Circulation Research, 66 (1990), pp. 1023–1039.
[35] J. Obergassel, I. Cazzoli, S. Guarguagli, V. D. Nagarajan, and S. Ernst, *Noninvasive* 3d *mapping and ablation of epicardial premature ventricular contractions from the endocardial aspect of the left atrial appendage*, JACC Case Reports, 2 (2020), pp. 1776–1780, https://doi.org/10.1016/j.jaccas.2020.07.045.
[36] H. S. Oster, B. Taccardi, R. L. Lux, P. R. Ershler, and Y. Rudy, *Noninvasive electrocardiographic imaging: Reconstruction of epicardial potentials, electrograms, and isochrones and localization of single and multiple electrocardiac events*, Circulation, 96 (1997), pp. 1012–1024.
[37] L. E. Payne, *Improperly Posed Problems in Partial Differential Equations*, SIAM, Philadelphia, 1975.
[38] R. Piersanti, P. C. Africa, M. Fedele, C. Vergara, L. Dedè, A. F. Corno, and A. Quarteroni, *Modeling cardiac muscle fibers in ventricular and atrial electrophysiology simulations*, Comput. Methods Appl. Mech. Engrg., 373 (2021), 113468.

[39] A. Quarteroni, A. Manzoni, and F. Negri, *Reduced Basis Methods for Partial Differential Equations: An Introduction*, UNITEXT 92, Springer, New York, 2015.
[40] M. Raissi, P. Perdikaris, and G. E. Karniadakis, *Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations*, J. Comput. Phys., 378 (2019), pp. 686–707.
[41] C. Ramanathan, R. N. Ghanem, P. Jia, K. Ryu, and Y. Rudy, *Noninvasive electrocardiographic imaging for cardiac electrophysiology and arrhythmia*, Nature Medicine, 10 (2004), pp. 422–428.
[42] C. Ramanathan and Y. Rudy, *Electrocardiographic imaging:* I. *Effect of torso inhomogeneities on body surface electrocardiographic potentials*, J. Cardiovascular Electrophysiology, 12 (2001), pp. 229–240.
[43] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Learning representations by back-propagating errors*, Nature, 323 (1986), pp. 533–536.
[44] R. Sameni, M. B. Shamsollahi, C. Jutten, and G. D. Clifford, *A nonlinear bayesian filtering framework for ECG denoising*, IEEE Trans. Biomedical Engineering, 54 (2007), pp. 2172–2185.
[45] O. Sayadi and M. B. Shamsollahi, *ECG denoising and compression using a modified extended Kalman filter structure*, IEEE Trans. Biomedical Engineering, 55 (2008), pp. 2240–2248.
[46] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, *Dropout: A simple way to prevent neural networks from overfitting*, J. Mach. Learn. Res., 15 (2014), pp. 1929–1958.
[47] J. Sundnes, G. T. Lines, X. Cai, B. F. Nielsen, K.-A. Mardal, and A. Tveito, *Computing the Electrical Activity in the Heart*, Vol. 1, Springer, New York, 2007.
[48] N. N. Tarkhanov, *The Cauchy Problem for Solutions of Elliptic Equations*, Mathematical Topics, 7, Vch Pub, 1995.
[49] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*, Winston, New York, 1977.
[50] L. Tung, *A Bi-Domain Model for Describing Ischemic Myocardial DC Potentials*, Ph.D. thesis, Massachusetts Institute of Technology, 1978.
[51] D. Wang, R. M. Kirby, and C. R. Johnson, *Resolution strategies for the finite-element-based solution of the ECG inverse problem*, IEEE Trans. Biomedical Engineering, 57 (2009), pp. 220–237.
[52] L. Wang, W. Wu, Y. Chen, and C. Liu, *An ADMM-net solution to inverse problem of electrocardiology*, in Proceedings of the 5th International Conference on Systems and Informatics, IEEE, 2018, pp. 565–569.
[53] Q. Wang, J. S. Hesthaven, and D. Ray, *Non-intrusive reduced order modeling of unsteady flows using artificial neural networks with application to a combustion problem*, J. Comput. Phys., 384 (2019), pp. 289–307.
[54] J. B. W. Webber, *A bi-symmetric log transformation for wide-range data*, Measurement Science Technology, 24 (2012), 027001.
[55] C. Wyndham, T. Smith, M. K. Meeran, R. Mammana, S. Levitsky, and K. Rosen, *Epicardial activation in patients with left bundle branch block*, Circulation, 61 (1980), pp. 696–703.
[56] C. R. Wyndham, M. K. Meeran, T. Smith, A. Saxena, R. M. Engelman, S. Levitsky, and K. Rosen, *Epicardial activation of the intact human heart without conduction defect*, Circulation, 59 (1979), pp. 161–168.
[57] N. Zemzemi, C. Dobrzynski, L. Bear, M. Potse, C. Dallet, Y. Coudière, R. Dubois, and J. Duchateau, *Effect of the torso conductivity heterogeneities on the ECGI inverse problem solution*, in Proceedings of the Computing in Cardiology Conference, IEEE, 2015, pp. 233–236.
[58] M. Zihlmann, D. Perekrestenko, and M. Tschannen, *Convolutional recurrent neural networks for electrocardiogram classification*, in Proceedings of the Computing in Cardiology Conference, IEEE, 2017, pp. 1–4.