POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Homework report

ENIGMA

Scientific computing tools for advanced mathematical modelling

Authors: Benedetta Bertesago, Leonardo Breda, Samuele Caldarini

Advisor: Prof. Stefano Pagani

Academic year: 2023-2024

---

0. Introduction

Atrial arrhythmias, including atrial fibrillation, are a major global health problem; the latter was estimated to affect 7.6 million people over 65 in the EU in 2016 [1].
The data used in this project comes from *electro-anatomical maps*, really helpful for clinicians to understand the movements of the cardiac tissue during electric impulses.

- **Electro**: data collected from intracardiac electrograms that help to understand the activation timing of the impulses and also the health status of the tissue through the signal amplitude
- **Anatomical**: represents the points on the surface of the ventricle, for simplicity only (x, y) were given, plus the activation times of each point in $\mathbb{R}^2$. These points are 20, collected for each patient, with a total of 100 patients.
- **Maps**: allow to record and catalogue both electrical activity and surface shape in order to display data more comprehensively.

Data are acquired using a mapping catheter with five splines, each containing four sensors that record the chamber's electrical activity at different locations.
The quantities of interest are: the curve for activation times (AT), and the associated conduction velocity (CV) for each point in the area under study.

1. 1st checkpoint

1.1. Mathematical formulation

The task is to adopt a fully data-driven approach to reconstruct AT and CV from sparse measurements on the entire domain $\mathcal{D} = \{(x, y) \in \mathbb{R}^2 \mid (x, y) \in (-1.5, 1.5) \times (-1.5, 1.5)\}$. The goal is finding a function $t(\mathbf{x})$ with $\mathbf{x} \in \mathcal{D}$ for each patient, such that, given some measurements $\{\tilde{x}_i, \tilde{t}_i\}$ $i = 1, \dots, 20$, the function satisfies:

- $t(\tilde{x}_i) = \tilde{t}_i$, i.e. interpolation constraint
- $t(\mathbf{x})$ is differentiable, for computing CV

A preliminary exploratory analysis of the samples in the specimen clarifies one aspect about the orientation of the sensors of the mapping catheter, that will be crucial in better understanding the shape of CV. Indeed, one of the five branches composing the structure is kept fixed and its angle put to 0 radians for all subjects. Instead the other arms have different orientations in space from sample to sample, allowing for more variability.

1.2. Method

The intended strategy to use is interpolation: from the twenty measurements, a curve interpolating all the points is reconstructed by means of radial basis functions (RBF) [2]. The reconstruction is performed sample-wisely,

obtaining 100 surfaces in $\mathcal{D}$ depicting AT and, deriving it, a map for atrial CV can be retrieved as well. Let $t_{AT}(\mathbf{x})$ denote the function to be reconstructed by means of RBF interpolation, it is given by:

$$t_{AT}(\mathbf{x}) = \sum_{i=1}^{N} w_i \cdot \phi(\|\mathbf{x} - \mathbf{x}_i\|_2) \tag{1}$$

with $\mathbf{x}_i \in \mathcal{D},\ i = 1, \dots, N = 20$ being the measurements, w_i the weights, and $\phi(\cdot)$ representing the kernel. After having selected **thin plate splines** as the appropriate kernel to get a differentiable estimate of the AT curve, so that $\phi(r) = r^2 log(r)$, CV can be numerically obtained using:

$$\text{CV} = \frac{1}{\|\nabla(t_{AT})\|_2 + \epsilon} \tag{2}$$

The command used to perform such interpolation allows for the tuning of a smoothing parameter λ. The choice of the latter is based upon the estimated CV curve retrieved from Equation 2. Indeed a value of CV greater than 150 cm/s is not physically attainable, therefore two non-negligible values for λ were set: $\lambda = 1/10$ and $1/2$. In this way the assumption of perfect interpolation of the data-points given by Equation 1 is relaxed. In order to still have a meaningful reconstruction of AT that accounts for the inclusion of the 20 measurements, a trade-off must be faced. Indeed, for each patient, a weighted average is considered to compute an efficient estimate of CV. Such average includes two terms: $w \cdot CV_{\lambda_1} + (1 - w) \cdot CV_{\lambda_2}$ where $CV_{\lambda_i}\ i = 1, 2$ is the conduction velocity computed with two different smoothing parameters following (2). For what concerns the weight w, it was set in order to account for the physical constraint and simultaneous smoothness of CV. It is indeed defined for the entire grid as follows:

$$w = w(x, y) = \frac{CV(x, y)}{\max_{x,y} CV(x, y)} \quad \forall\ x, y \in \mathcal{D} \tag{3}$$

and then each value is raised to the power of 1.4 and multiplied by 0.95 in order to not alter the estimates too much, but still give an additional smoothing effect to the sum.

1.3. Numerical results

After obtaining a reconstruction of the activation times evaluated at each point of a structured grid over $\mathcal{D}$, here is an approximation of AT and the corresponding conduction velocity (for the latter, the unit of measure on the z-axis is cm/s, Figure 1).

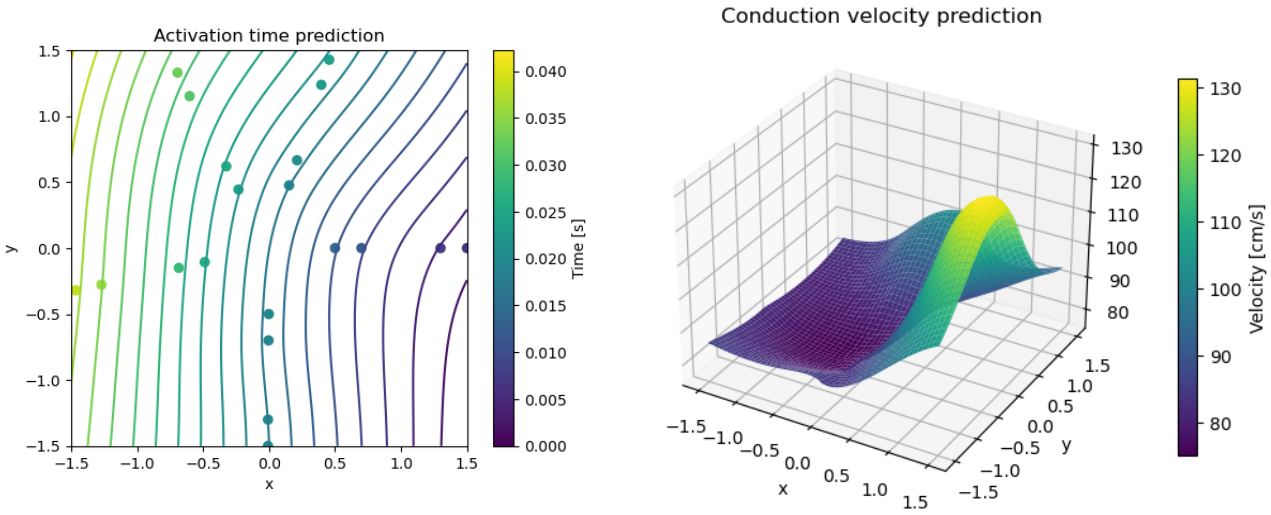

Figure 1: Plot of one patient's AT and CV reconstruction from twenty measurements

2. 2nd checkpoint

The second task is to determine an estimate of an unknown vector of parameters $\boldsymbol{\mu}$. Such a vector contains three quantities that are patient-specific, namely:

1. **Fiber angle** (μ_1): indicates the angle of propagation (setting the horizontal to 0 radians) of the electric shock-wave. It's bounded in $[-\frac{\pi}{10}; \frac{\pi}{10}]$.
2. **Anisotropy ratio** (μ_2): is a quantification of the sphericity of the wavefront. A value close to 1 indicates an almost spherical front, while a value greater than 1 states that the wavefront assumes a more ellipsoidal shape. It's bounded in $[1; 9]$.
3. **Starting point** ($(1.5; \mu_3)$): can be referred to as the source of the shock-wave within the domain $\mathcal{D}$. It's bounded in $(-1.5; 1.5)$.

2.1. Mathematical formulation

The role of $\boldsymbol{\mu}$ is better understood by stating the physical law that the model is subjected to, here:

$$\begin{cases} 100\sqrt{\nabla t_{AT}(x,y)\ \mathbf{D}\ \nabla t_{AT}(x,y)} = 1, & (x,y) \in \mathcal{D} \quad (4a) \\ t_{AT}(1.5, \mu_3) = 0, & \mu_3 \in (-1.5, 1.5) \quad (4b) \end{cases}$$

with D being the following tensor:

$$\mathbf{D} = \frac{1}{\mu_2} \begin{bmatrix} \cos(\mu_1) \\ \sin(\mu_1) \end{bmatrix} \otimes \begin{bmatrix} \cos(\mu_1) \\ \sin(\mu_1) \end{bmatrix} + \begin{bmatrix} \cos(\mu_1 - \pi/2) \\ \sin(\mu_1 - \pi/2) \end{bmatrix} \otimes \begin{bmatrix} \cos(\mu_1 - \pi/2) \\ \sin(\mu_1 - \pi/2) \end{bmatrix}$$

Equation 4a is the Eikonal equation, typically used in signal propagation problems. It puts in relation a generic field (in this case AT) with its gradient (here corresponding to CV) and it helps to understand the movement on the shock-wave inside the medium or on its surface. The fact that CV is not computed for this step is because it is set equal to a constant field with value 100 cm/s all over the domain.

To determine $\boldsymbol{\mu}$ we have to solve a minimization problem, whose loss function is:

$$J(\boldsymbol{\mu}) = \sum_{i=1}^{20} \|t_i - F(\mathbf{x}_i; \boldsymbol{\mu})\|^2 \quad (5)$$

It takes into account also $\boldsymbol{\mu}$, so not only a problem of interpolation is faced, but also of parameter estimation. For what concerns the above $\boldsymbol{\mu}$, a guess for its initial value is provided to the forward map $F(\cdot\,;\cdot)$ that will try to compute a proper estimate. Alongside this procedure, the whole curve of the activation times (AT) needs to be reconstructed as well. Indeed this follows exactly section 1.1 except for the computation of CV which is useless for the current aim, and so it is neglected.

2.2. Method

Regarding the initial guess for $\boldsymbol{\mu}$ it was so provided: $[0.01; 3; y_0]^\top$, where y_0 is obtained minimizing AT from Equation 1 on the right-boundary of $\mathcal{D}$ ($y_0 = \min_y AT(1.5, y) \quad y \in \mathcal{D}$).
Given the presence of a physical constraint to be satisfied, a purely data-driven model would under-perform in this scenario. So a Physics-Informed Neural-Network (PINN) is adopted.
For what concerns the structure of the final network used, it is so composed:

- 64 neurons of *tanh* activation function
- 64 neurons of *mish*[1] activation function
- 64 neurons of *tanh* activation function
- final output layer

For each inner dense layer a L^2 regularizer is added in order to have, during the training, an additional term summed to the loss function whose aim is to penalize large weights and stimulate the network to follow simpler paths.
Lastly the collocation points used to fit the PINN output were sampled from a Sobol sequence on the domain $\mathcal{D}$ instead of the usual *latin hypercube* and in total they were $2^{14} \approx 16.000 = N_{cp}$.

[1] mish: is obtained as $x \cdot sigmoid(x)$, is similar to a leaky ReLU in the shape and so is differentiable at 0

The formulation of the loss function is obtained by extending the general writing in Equation 5. Indeed it has now a larger structure with additional terms, some of them not deriving from the AT reconstruction, but essential to fulfill the physical constraints:

$$J(\boldsymbol{\mu}) = \sum_{i=1}^{20} \|t_i - PINN(\mathbf{x}_i; \boldsymbol{\mu})\|^2 + R_P^2 + R_B^2 + R_I^2 + \lambda \|w\|_2^2 \tag{6}$$

where:

- $R_P = \frac{1}{N_{cp}} \sum_{j=1}^{N_{cp}} \left| 100\sqrt{\nabla PINN\,(x_j;\boldsymbol{\mu})\ \cdot \mathbf{D}\ \cdot \nabla PINN\,(x_j;\boldsymbol{\mu})} \ - \ 1 \right|$ is the term that ensures the Eikonal equation is satisfied across the whole domain $\mathcal{D}$ as in Equation 4a
- $R_B = \max(\text{lower bound} - \boldsymbol{\mu}, 0) + \max(\boldsymbol{\mu} - \text{upper bound}, 0)$ accounts for the boundaries of each component of $\boldsymbol{\mu}$. In order to have meaningful estimates, they must lie in their own support
- $R_I = |PINN(1.5; \mu_3)|$ is added in order to have the initial condition in Equation 4b met
- the last term accounts for the added L^2 regularization on the weights w of the network

2.3. Numerical results

Hyperparameters are tuned according to gold standard values provided, obtained from the anisotropic fast marching method (AFMM) [3].
Here are some results.

	True value	Estimate
Fibre angle (rad)	0.0324	0.0974
Anisotropy ratio	6.6297	6.8531
Starting point (mm)	-0.1749	-0.2701

Table 1: Results for patient #1

All the values can be checked visually from the plot of AT: indeed the angle is almost null, the shape of the isochrones (near the source) is ellipsoidal and the starting point of the shock-wave is right below 0.

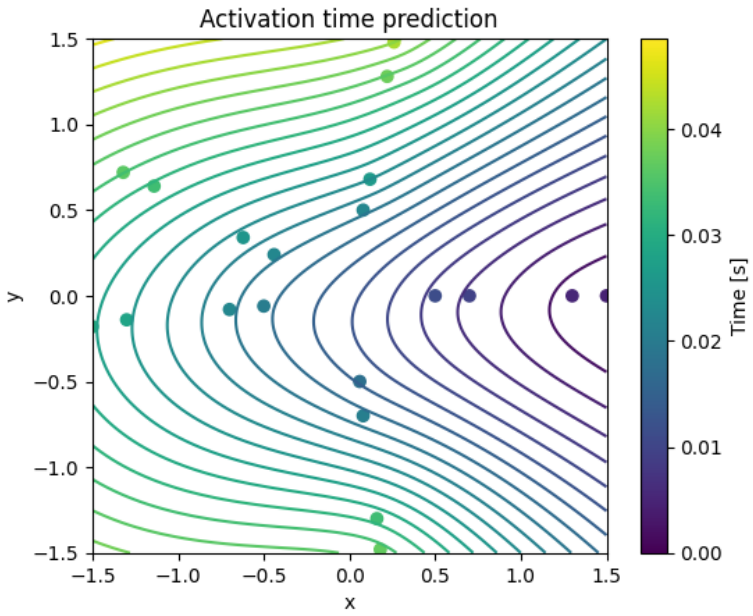

Figure 2: Contour lines of the AT reconstruction for patient #1

3. 3rd checkpoint

3.1. Mathematical formulation

The final checkpoint is a natural extension of the previous task. Indeed, Equation 4a can be fully exploited by substituting the constant value of 100 cm/s previously set and introducing a general velocity field $CV(x, y)$. The updated equation is:

$$CV(x,y)\sqrt{\nabla t_{AT}(x,y)\ \mathbf{D}\ \nabla t_{AT}(x,y)} = 1, \qquad (x,y) \in \mathcal{D} \tag{7}$$

with $\mathbf{D}$ now including a fixed estimate of $\boldsymbol{\mu}$ for all the patients:

$$D = \frac{1}{\hat{\mu}_2}\begin{bmatrix}\cos(\hat{\mu}_1)\\ \sin(\hat{\mu}_1)\end{bmatrix} \otimes \begin{bmatrix}\cos(\hat{\mu}_1)\\ \sin(\hat{\mu}_1)\end{bmatrix} + \begin{bmatrix}\cos(\hat{\mu}_1 - \pi/2)\\ \sin(\hat{\mu}_1 - \pi/2)\end{bmatrix} \otimes \begin{bmatrix}\cos(\hat{\mu}_1 - \pi/2)\\ \sin(\hat{\mu}_1 - \pi/2)\end{bmatrix}$$

The main goal is to reconstruct the field $CV(\cdot, \cdot)$ for a patient, given its twenty measurements. For a hundred patients the estimate of the field is provided. In this way is possible to train properly the structure responsible for the reconstruction of the field of future patients.

3.2. Method

The aim is to build a surrogate model that given the measurements of a new patient and the scores coming from Principal Component Analysis (PCA) on $CV(x, y)$ of available data, is able to reconstruct a compressed representation of the field. Afterwards, the inverse transformation is applied to obtain the final values for $CV(x, y),\ \forall\ (x, y) \in \mathcal{D}$.
As a first sub-task, an **estimate of $\hat{\boldsymbol{\mu}}$** is computed according to section 3.1. The value of $\boldsymbol{\mu} = \hat{\boldsymbol{\mu}}$ is now fixed and shared among every patient. To have a more reliable estimate of such vector, the $\boldsymbol{\mu}$'s of the two patients having a more similar behaviour of the field to the constant value of 100 cm/s ($CV(x, y) \approx 100\ \forall\ (x, y) \in \mathcal{D}$) are averaged together. The criterion for similarity was established by computing the Frobenius norm of the difference of the field and the reference value of 100 cm/s:

$$\|CV(x,y) - 100 \cdot \mathcal{I}_{151\times151}\|_F$$

where $\mathcal{I}_{151\times151}$ represents a matrix full of ones i.e. the constant field.
The reconstructing procedure is preceded by a pre-processing step. It is more practical, indeed, to handle a portion of the entire field $CV(x, y)$ instead of its whole structure which turns out to be of dimension $151 \times 151 = 22801$ for a single patient. Hence a problem of dimensionality reduction must be faced. The advantages of introducing it are twofold: in the first place it allows us to decrease the computational effort, secondly, since the number of hyperparameters is consequently smaller, it is less prone to overfit and so the results are more reliable.
The technique chosen is the **Principal Component Analysis**, performed on scaled data. This method transforms data into a new smaller system of coordinates, reducing the number of variables while retaining most of the information. Looking at the cumulative variance explained, the first five PCs were considered enough, since they are able to capture the 90% of the total variability.

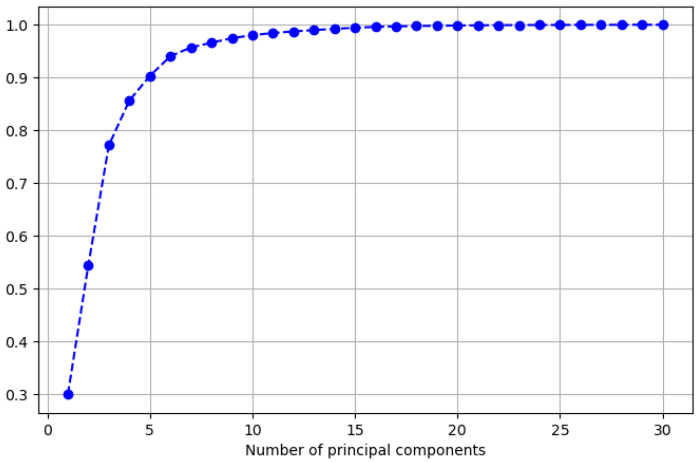

Figure 3: Cumulative explained variance

As last step, let us define a **surrogate model** for predicting activation times for a new patient, given the compressed representation of the field of 100 patients and the twenty measurements of this new one. This prediction is used in the loss function of a minimization routine that is able to find the best values of the five PCs, later employed in the inverse PCA transformation, that will output the whole speed field for a new patient. The surrogate model chosen relies on **Polynomial Chaos Expansion** (PCE) technique: a tool that represents the output of a complex model as a series of polynomial terms. Mathematically, given a random variable $Y(\boldsymbol{\xi})$, PCE represents it as $Y(\boldsymbol{\xi}) \approx \sum_{i=0}^{N} c_i \Psi_i(\boldsymbol{\xi})$, where c_i are expansion coefficients and Ψ_i are orthogonal polynomials. In particular, the following surrogate model is defined to predict the activation times in the twenty locations (x, y) for a given patient:

$$\hat{F}(t) = \sum_{k} c_k \psi_k(x, y, PC_1, PC_2, PC_3, PC_4, PC_5) \tag{8}$$

where k represents the degrees of polynomials. For the coordinates of given measurements uniform distributions were used in range $(-1.5, 1.5)$, covering the entire domain.
By analyzing the distribution of the scores of each PC, a proper distribution in closed form is used to approximate each of them. The following picture shows them:

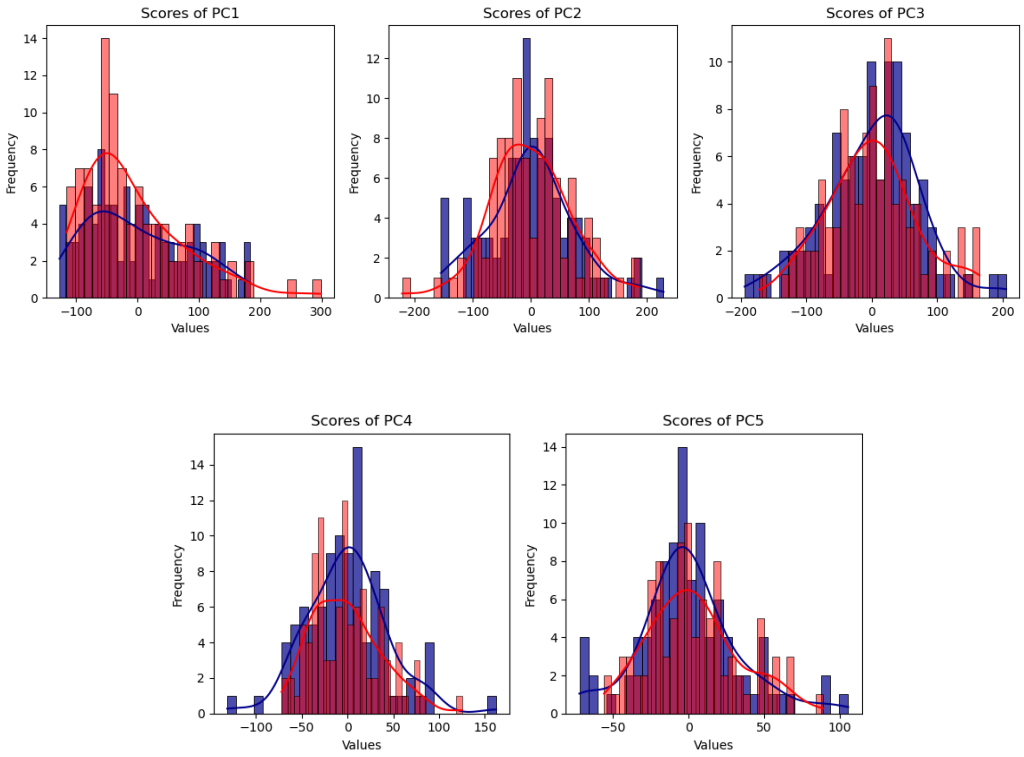

Figure 4: Plot of the PC: in blue the true distribution and in red the approximation

The approximations chosen are Gamma distributions for 1^{st} and 5^{th} PCs, while 2^{nd}, 3^{rd} and 4^{th} PCs were selected to be Gaussians distributions. Means and variances were tuned properly in order to match the sample means and variances as well.

In Figure 5 one can notice on the right the approximated distributions employed for PCE matching the corresponding on the left. Figure 6 is showing a DD plot, a tool used to visually asses whether two distributions are the same or not: if the shape in the middle is circular or almost-circular without spikes or clusters, the two distributions can be considered the same, as in this case.

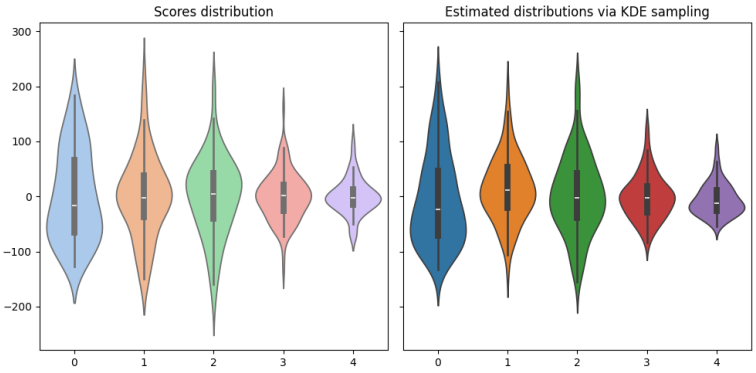

Figure 5: True and approximated distribution for chosen PCs

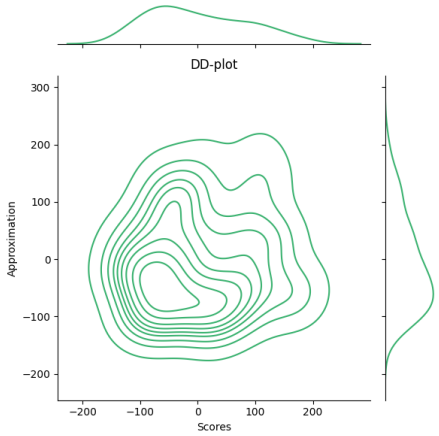

Figure 6: DD plot for PC_1

After having set the appropriate distributions for the sampling phase of the above mentioned weights, the function *scipy.optimize.minimize* was employed with *Powell* minimizer to take care of the loss function measuring the discrepancy between the true activation time and the one associated to the PCE reconstruction.
This procedure was done twice: the first with all the polynomial degrees for PCE set equal to 2, while the other with cubic basis functions. Then the results were combined with a weighted average in favour of the quadratic basis because of the parabolic shapes of the fields $C(x, y)$ given.

3.3. Numerical results

It is possible to visualize the true and reconstructed speed field as well as the difference between them, here an example.

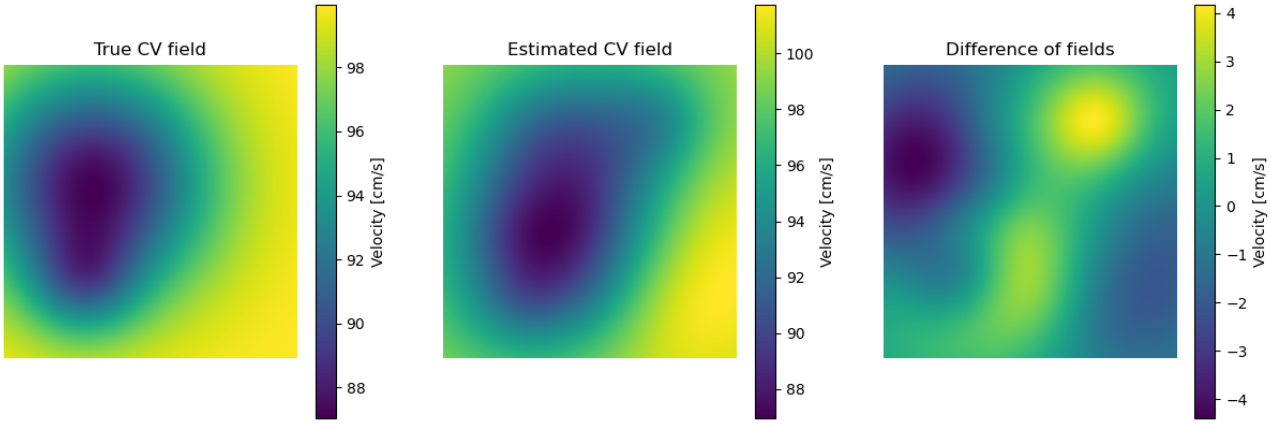

Figure 7: Plot of the results for a patient

In Figure 7 the difference was not exceeding 5 cm/s in absolute value. However there were cases where the reconstruction was not so good at approximating the true fields, but overall:

- Median RMSE = 4.89
- Mean RMSE = 5.13

just by computing the root mean square error on the entire grid $\mathcal{D}$ for the difference of the two fields.

4. Conclusions

What can be said about the three tasks is the following:

- RBF interpolation is an efficient and flexible tool for curve reconstruction given only sparse measurements thanks to *ad hoc* choices of kernels and smoothing terms. The reconstruction of conduction velocity appears more delicate since it is computed with Equation 2 that displays AT at the denominator, to avoid numerical errors, an ϵ was added to ensure numerical stability.
  In order to stay within the boundaries dictated from the physical phenomenon, the decision to add

smoothing terms in the time interpolation was taken. Relaxing the assumption of perfect interpolation, the results were still precise enough to affirm that this technique is effective.

- Second checkpoint introduces physical knowledge that was included using PINNs. This tool is very powerful, but has a significant trade-off between efficiency and accuracy: for obtaining precise results of quantities of interest, a huge number of collocation points have been used, leading to a very long execution time. Another time-consuming task is the fine-tuning of all the other hyperparameters, such as the number of neurons per layer and the choice of suitable activation functions.
- Last checkpoint requires the reconstruction of a vector field in all the domain $\mathcal{D}$, namely 22801 points. Facing it with a model that directly reconstructs all the values without any pre-processing is difficult and useless at the same time. Thus we proceed firstly by reducing the dimensionality, performing PCA on provided data. After this we chose PCE as surrogate model for predicting activation times given the compressed representation of the field and the twenty measurements. This prediction is then used in the loss function of a minimization routine able to find the best values of the five PCs. PCE performs a linear regression and it's indeed the main drawback: this could make it a less precise instrument to use, but a lot quicker and interpretable with respect to other approaches, e.g. neural networks.

References

[1] Antonio Di Carlo, Leonardo Bellino, Domenico Consoli, Fabio Mori, Augusto Zaninelli, Marzia Baldereschi, Alessandro Cattarinussi, Maria Grazia D'Alfonso, Chiara Gradia, Bruno Sgherzi, et al. Prevalence of atrial fibrillation in the italian elderly population and projections from 2020 to 2060 for italy and the european union: the fai project. *Ep Europace*, 21(10):1468–1475, 2019.

[2] M Masé and F Ravelli. Automatic reconstruction of activation and velocity maps from electro-anatomic data by radial basis functions. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pages 2608–2611. IEEE, 2010.

[3] Maxime Sermesant, Yves Coudière, Valérie Moreau-Villéger, Kawal S Rhode, Derek LG Hill, and RS Razavi. A fast-marching approach to cardiac electrophysiology simulation for xmr interventional imaging. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2005: 8th International Conference, Palm Springs, CA, USA, October 26-29, 2005, Proceedings, Part II 8*, pages 607–615. Springer, 2005.