MOX–Report No. 19/2013

Blood flow velocity field estimation via spatial regression with PDE penalization

Azzimonti, L.; Sangalli, L.M.; Secchi, P.; Domanin, M.; Nobile, F.

MOX, Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox@mate.polimi.it http://mox.polimi.it

Blood flow velocity field estimation via spatial regression with PDE penalization

Laura Azzimonti[1] Laura M. Sangalli[1] Piercesare Secchi[1]
Maurizio Domanin[2,3] Fabio Nobile[1,4]

April 30, 2013

[1] MOX - Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, Italy (e-mail: laura.sangalli@polimi.it)
[2] U.O. di Chirurgia Vascolare Fondazione I.R.C.C.S. Ca' Granda Ospedale Maggiore Policlinico, Milano
[3] Dipartimento di Scienze Chirurgiche Specialistiche, Università di Milano
[4] MATHICSE-CSQI, École Polytechnique Fédérale de Lausanne, Switzerland

Abstract

We propose an innovative method for the accurate estimation of surfaces and spatial fields when a prior knowledge on the phenomenon under study is available. The prior knowledge included in the model derives from physics, physiology or mechanics of the problem at hand, and is formalized in terms of a partial differential equation governing the phenomenon behavior, as well as conditions that the phenomenon has to satisfy at the boundary of the problem domain. The proposed models exploit advanced scientific computing techniques and specifically make use of the Finite Element method. The estimators have a typical penalized regression form and the usual inferential tools are derived. Both the pointwise and the areal data frameworks are considered. The driving application concerns the estimation of the blood-flow velocity field in a section of a carotid artery, using data provided by echo-color doppler; this applied problem arises within a research project that aims at studying atherosclerosis pathogenesis.

Keywords: functional data analysis, spatial data analysis, object-oriented data analysis, penalized regression, Finite Elements.

1 Introduction

In this work we propose a novel non-parametric regression technique for surface and spatial field estimation, able to include a prior knowledge on the shape of the surface or spatial field and to comply with complex conditions at the boundary of the problem domain. The motivating applied problem concerns the estimation of the blood-flow velocity field on a cross-section of an artery, using data provided by echo-color doppler acquisitions. This study is carried out within the project *MAthematichs for CARotid ENdarterectomy @ MOX*[1], that gathers researchers in statistics, numerical analysis and computer sciences and medical doctors in cardiac surgery, with the aim of investigating the pathogenesis of atherosclerosis in human carotids. The project intends specifically to study the role of blood fluid-dynamics and vessel morphology on the formation process and histological properties of atherosclerotic plaques. Interactions between the hemodynamics and atherosclerotic plaques have been highlighted for instance in [10] via numerical simulations of the blood flow on real patient-specific vessel morphologies.

The data collected within the project include: Echo-Color Doppler (ECD) measurements of blood flow at a cross-section of the common carotid artery, 2 cm before the carotid bifurcation, for patients affected by high grade stenosis (>70%) in the internal carotid artery; the reconstruction of the shape of this cross-section obtained via segmentation of Magnetic Resonance Imaging (MRI) data. The first phase of the project requires the estimation, starting from these data, of the blood-flow velocity fields in the considered carotid section. These estimates are first of all of interest to the medical doctors, as they highlight relevant features of the blood flow, such as the eccentricity and the asymmetry of the flow or the reversion of the fluxes, which could have an impact on the pathology. Moreover, they will enable a population study that explores quantitatively the relationship between the blood-flow and the atherosclerosis. Finally, the estimated blood velocity fields will also be used as patient-specific and physiological inflow conditions for hemodynamics simulations, that in turns aim at further enhancing the knowledge on this relationship.

Carotid Echo-Color Doppler (ECD) is a medical imaging procedure that uses reflected ultrasound waves to create images of an artery and to measure the velocity of blood cells in some locations within the artery. This technique does not require the use of contrast media or ionizing radiation and has relative low costs. Thanks to this complete non-invasivity and also to the short acquisition time required, ECD scans are largely used in clinics, even though they provide a less rich and noisier information than other diagnostic devices, such as Phase Contrast Magnetic Resonance Imaging. The left panel of Figure 1 shows one of the ECD images used in the study. The ultrasounds image in the upper part of the figure represents the longitudinal section of the vessel. It also shows by a small gray box the position of the beam where blood particle velocities, in the longitudinal direction of the vessel, are measured; the dimension of the box relates to the dimension of the beam. In the case considered in this picture, the acquisition beam is located in the center of the considered cross-section of the artery. The lower part of the ECD image is a graphical display of the acquired velocity signal during the time lapse of about four heart beats. This signal represents the histogram of the measured velocities, evolving in time. More precisely, the x-axis represents time and the y-axis represents velocity classes; for any fixed time, the gray-scaled intensity of pixels is proportional to the number of blood-cells in the beam moving at a certain velocity. For the purpose of this work, we shall consider a fixed time instant corresponding

[1]The *MACAREN@MOX* project involves clinicians from Ca' Granda Ospedale Maggiore Policlinico in Milano, statisticians, numerical analysis and image processing scientists from MOX Laboratory for Modeling and Scientific Computing, Politecnico di Milano, and numerical analysis scientists from Università degli studi di Bergamo and École Polytechnique Fédérale de Lausanne. Principal Investigator of the project is Dr. Christian Vergara.

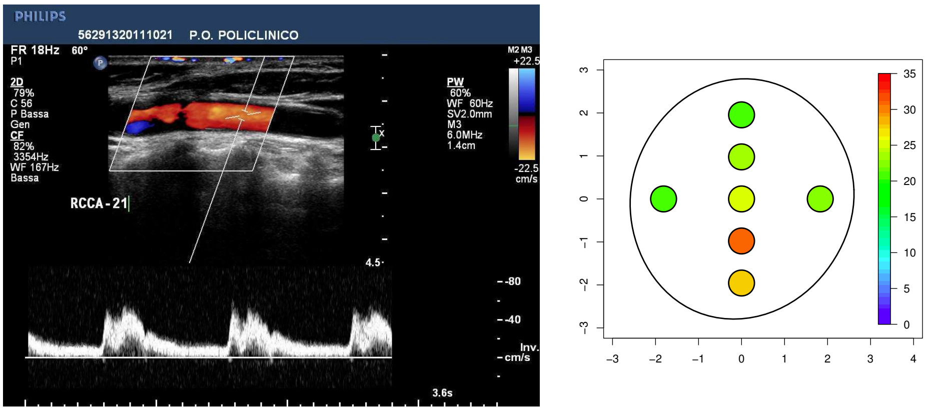

Figure 1: Left panel: ECD image corresponding to the central point of the carotid section located 2 cm before the carotid bifurcation. Right panel: MRI reconstruction of the cross-section of the carotid artery located 2 cm before the bifurcation; cross-shaped pattern of observations with each beam colored according to the mean blood-velocity measured on the beam at systolic peak time.

to the systolic peak, which is of crucial clinical interest.

The right panel of Figure 1 shows the reconstruction from MRI data of the considered cross-section of the carotid artery; it also displays the spatial location of the beams inspected in the ECD scan. In particular, during the ECD scan 7 beams are considered, located in a cross-shaped pattern; this unusual pattern is a compromise decided together with clinicians in order to obtain as many observations as possible in the short time dedicated to the acquisition. In the figure each beam is colored according to the value of the mean velocity registered within the beam at the fixed time instant considered, the systolic peak.

In this applied problem there are important conditions at the boundary of the problem domain, i.e., specifically, at the wall of the carotid cross-section represented in Figure 1, Right. The physics of the problem implies in fact that blood-flow velocity is zero at the arterial wall, due to the friction between blood cells and arterial wall; these are the so-called no-slip boundary conditions. Classical methods for surface estimation as thin-plate splines, tensor product splines, kernel smoothing, wavelet-based smoothing and kriging, do not naturally include information on the shape of the domain and on the value of the surface at the boundary, although it is possible to enforce such boundary conditions for example with binning. Recently, some methods have been proposed where the shape of the domain and the boundary conditions are instead directly specified in the model. For instance, Finite Element L-splines described in [15] account explicitly for the shape of the domain, efficiently dealing with irregular shaped domains; soap-film smoothing (SOAP), described in [19], considers both the shape of the domain and some common types of boundary conditions; Spatial Spline Regression (SSR), presented in [16], extends [15] and includes general boundary conditions. The methods in [15], [19] and [16] are penalized regression methods with a roughness term involving the Laplacian of the field, the Laplacian being a simple form of partial differential operator that provides a measure of the local curvature of the field. Although being able to account for the shape of the domain and to comply with the required boundary conditions, these methods do not provide physiological estimates of the velocity field. Figure 2 shows for example the velocity field estimated using SSR. The penalization of a measure of the local curvature of the field oversmooths and flattens the field toward a plane in those regions of the domain where no observations are available; the resulting estimated velocity field has thus rhomboidal isolines, which are certainly non-physiological.

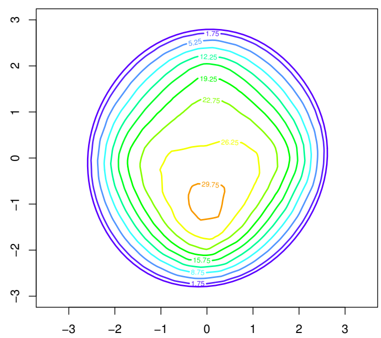

Figure 2: Estimate of the blood-flow velocity field on the carotid section using standard SSR.

On the other hand, we have prior knowledge on the phenomenon under study that could be exploited to derive accurate physiological estimates. There is in fact a vast literature devoted to the study of fluid dynamics and hemodynamics, see for example [5] and references therein. For what serves our purpose, it suffices to know that the theoretical solution of a stationary velocity field in a straight pipe with circular section has a parabolic profile. In our application, during the systolic phase, we hence expect to obtain a velocity field similar in shape to a parabolic profile, with isolines resembling circles. Notice that the real blood velocity field is not perfectly parabolic due to the curvature of the artery, the non-stationarity of the blood flow and the imperfect circularity of the artery section. For this reason, imposing a parametric model that forces a parabolic estimate would not be appropriate; such model would for instance completely miss asymmetries and eccentricities of the flow. Nevertheless, this prior information concerning the shape of the field, which can be conveniently translated into a Partial Differential Equation (PDE), could be incorporated in a non-parametric model, along with the desired boundary conditions.

In this work, extending [15] and [16], we propose a non-parametric model that includes prior information on the phenomenon under study, coming for instance from the physics, physiology, mechanics or chemistry of the problem, and formalized in terms of a governing PDE. Specifically, Spatial Regression with PDE penalization (SR-PDE) features a roughness term that involves, instead of the simple Laplacian, a more general PDE modeling the phenomenon. The applicability of the proposed method is by no way restricted to the problem here considered; PDEs are in fact commonly used to describe phenomena behavior in many fields of engineering and sciences, including bio-sciences, geo-sciences and physical sciences. It should be noticed that many methods for surface and spatial field estimation, besides the already cited SSR and SOAP, use roughness penalties involving some simple form of PDEs. A classical example is given by thin-plate-splines, while a recent proposal is offered for instance by [6]. Our work has also strong connections with the framework introduced by [8] that is based on a stochastic PDE. The novelty of the proposed SR-PDE models with respect to these methods is that the PDE is used to model the space variation of the phenomenon, using problem-specific information. Moreover SR-PDE allows for important modeling flexibility, accounting for space anisotropy and non-stationarity in a straightforward way, as well as unidirectional smoothing effects.

Likewise in [15] and [16], SR-PDE exploits advanced numerical analysis techniques and, specifically, it makes use of the Finite Element Method, which provides a basis for piecewise polynomial surfaces. The resulting estimators have a typical penalized regression form, they are linear in the observed data values and classical inferential tools can be derived. The proposed method is currently implemented in `R` [13] and in FreeFem++ [11].

The paper is organized as follows. Section 2 introduces SR-PDE for pointwise observations. Section 3 extends the models to the case of areal data, which is of interest in many applications, including the analysis of ECD measurements here considered. Section 4 describes the Finite Element solution to the estimation problem and derives the inferential properties of the estimators. Section 5 deals with general boundary conditions. In Section 6, SR-PDE is compared to standard SSR and to SOAP in different simulation settings, with data distributed uniformly on the domain or only on some subregions, showing that the inclusion of the prior knowledge on the phenomenon behavior improves significantly the estimates. In Section 7 the application within the MACAREN@MOX project is presented: details on the ECD acquisitions are given and the results obtained with SR-PDE are shown. Section 8 outlines future research directions.

2 Model for pointwise data

Consider a bounded and regular domain $\Omega \subset \mathbb{R}^2$, whose boundary $\partial\Omega$ is a curve of class C^2, and n observations z_i, for $i = 1, \ldots, n$, located at points $\mathbf{p}_i = (x_i, y_i) \in \Omega$. Assume the model

$$z_i = f_0(\mathbf{p}_i) + \epsilon_i \tag{1}$$

where ϵ_i, $i = 1, \ldots, n$, are independent errors with zero mean and constant variance σ^2, and $f_0 : \Omega \to \mathbb{R}$ is the surface or spatial field to be estimated. In our application, Ω will be the carotid cross-section of interest, the observations z_i will represent the blood particles velocities measured by ECD in the longitudinal direction of the artery (i.e., in the orthogonal direction to Ω) and the surface f_0 will represent the longitudinal velocity field on the carotid cross-section.

Assume that problem specific prior information is available, that can be described in terms of a PDE, $Lf_0 = u$, modeling to some extent the phenomenon under study; moreover, prior knowledge could also concern possible conditions that f_0 has to satisfy at the boundary $\partial\Omega$ of the problem domain. Generalizing the models in [15] and [16], we propose to estimate f_0 by minimizing the penalized sum-of-square-error functional

$$J(f) = \sum_{i=1}^{n} (f(\mathbf{p}_i) - z_i)^2 + \lambda \int_{\Omega} \big(Lf(\mathbf{p}) - u(\mathbf{p})\big)^2 d\mathbf{p} \tag{2}$$

with respect to $f \in V$, where V is the space of functions in $L^2(\Omega)$ with first and second derivatives in $L^2(\Omega)$, that satisfy the required boundary conditions (b.c.). The penalized error functional hence trades off a data fitting criterion, the sum-of-square-error, and a model fitting criterion, that penalizes departures from a PDE problem-specific description of the phenomenon. In particular, we consider here phenomena that are well described in terms of linear second order elliptic operators L (with smooth and bounded parameters) and forcing term $u \in L^2(\Omega)$ that can be either $u = 0$, homogeneous case, or $u \neq 0$, non-homogeneous case. The operator L can include second order differential operators as the divergence of the gradient ($\mathrm{div}\nabla f$), first order differential operators as the gradient (∇f) and also the identity (f); the general form that we consider is

$$Lf = -\mathrm{div}(\mathbf{K}\nabla f) + \mathbf{b} \cdot \nabla f + cf \tag{3}$$

where the symmetric and positive definite matrix $\mathbf{K} \in \mathbb{R}^{2\times 2}$ is the diffusion tensor, $\mathbf{b} \in \mathbb{R}^2$ is the transport vector and $c \geq 0$ is the reaction term. Setting $\mathbf{K} = \mathbf{I}$, $\mathbf{b} = \mathbf{0}$, $c = 0$ and $u = 0$ we obtain the special case described in [15] and [16], where the Laplacian Δf is penalized, thus controlling the local curvature of f.

The three terms that compose the general second order operator (3) provide different smoothing effects. The diffusion term $-\mathrm{div}(\mathbf{K}\nabla f)$ induces a smoothing in all the directions; if the diffusion matrix $\mathbf{K}$ is a multiple of the identity the diffusion term has an isotropic smoothing effect, otherwise it implies an anisotropic smoothing with a preferential direction that corresponds to the first eigenvector of the diffusion tensor $\mathbf{K}$. The degree of anisotropy induced by the diffusion tensor $\mathbf{K}$ is controlled by the ratio between its first and second eigenvalue. It's possible to visualize the diffusion term as the quadratic form in $\mathbb{R}^2$ induced by the tensor $\mathbf{K}^{-1}$. On the contrary the transport term $\mathbf{b} \cdot \nabla f$ induces a smoothing only in the direction specified by the transport vector $\mathbf{b}$. Finally, the reaction term cf has instead a shrinkage effect, since penalization of the L^2 norm of f induces a shrinkage of the surface to zero.

The parameters of the PDE can be space-varying on Ω; i.e., $\mathbf{K} = \mathbf{K}(x, y)$, $\mathbf{b} = \mathbf{b}(x, y)$ and $c = c(x, y)$. This feature is fundamental to translate the a priori information on the phenomenon. For instance, in the blood flow velocity application, the problem specific prior information can be described via a differential operator that includes: a space varying anisotropic diffusion tensor that smooths the observations in the tangential direction of concentric circles (see Figure 3 Left); a transport field that smooths the observations in the radial direction, from the center of the section to the boundary (see Figure 3 Right). The reaction term is instead not required in this application. Notice that the space-varying parameters need to satisfy some regularity conditions to ensure that the estimation problem is well-posed, see [4] for details. The functional $J(f)$ is in fact well defined if $f \in V$ since $V \subset H^2(\Omega) \subset C(\bar{\Omega})$ if $\Omega \subset \mathbb{R}^2$ and the misfit of the PDE is square integrable.

We can impose different types of boundary conditions, homogeneous or not, that involve the evaluation of the function and/or its first derivative at the boundary, allowing for a complex modeling of the behavior of the surface at the boundary $\partial\Omega$

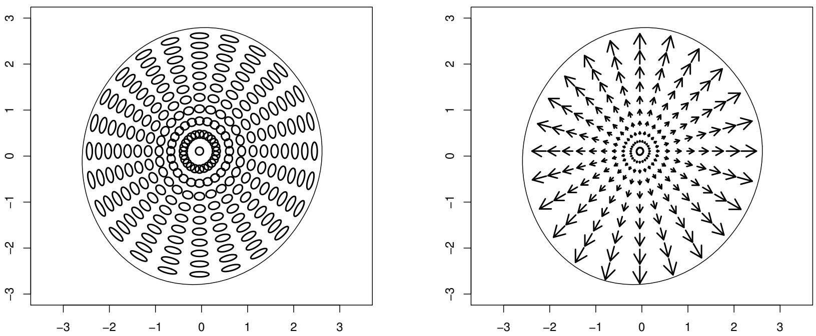

Figure 3: Left: diffusion tensor field $\mathbf{K}$, used in the velocity field application, that smooths the observations in the tangential direction of concentric circles. Right: transport field $\mathbf{b}$, used in the velocity field application, that smooths the observations in the radial direction, from the center of the section to the boundary.

of the domain. For ease of notation we consider in the following the simple case of homogeneous Dirichlet b.c., which involve the value of the function at the boundary, clamping it to zero, i.e., $f|_{\partial\Omega} = 0$. These boundary conditions correspond to the physiological no-slip conditions needed in the ECD application; the blood cells have in fact zero longitudinal velocity near the arterial wall due to friction between the particles and the arterial wall. In Section A.2 we extend all the results presented in this section to the case of more general non-homogeneous boundary conditions that can also involve first derivatives. In this work the boundary conditions are directly included in the space V; in the case of Dirichlet homogeneous b.c., V is the space of functions in $L^2(\Omega)$ with first and second derivatives in $L^2(\Omega)$ and zero value at the boundary $\partial\Omega$.

To lighten the notation, surface integrals will be written without the integration variable $\mathbf{p}$; unless differently specified, the integrals are computed with respect of the Lebesgue measure, i.e., $\int_D q = \int_D q(\mathbf{p})d\mathbf{p}$, for any $D \subseteq \mathbb{R}$ and integrable function q.

All the results presented can also be extended to include space-varying covariate information, following the semi-parametric approach described in [16].

2.1 Solution to the estimation problem

The estimation problem can be formulated as follows.

Problem 1. *Find $\hat{f} \in V$ such that*

$$\hat{f} = \underset{f \in V}{\operatorname{argmin}}\, J(f).$$

Existence and uniqueness of the surface estimator $\hat{f}$ are established in the following proposition.

Proposition 1. *Under suitable regularity conditions for L, the solution of Problem 1 exists and is unique. The surface estimator $\hat{f}$ is obtained by solving:*

$$\begin{cases} L\hat{f} = u + \hat{g} & \text{in } \Omega \\ \hat{f} = 0 & \text{on } \partial\Omega \end{cases} \qquad \begin{cases} L^*\hat{g} = -\frac{1}{\lambda}\sum_{i=1}^{n}(\hat{f} - z_i)\delta_{\mathbf{p}_i} & \text{in } \Omega \\ \hat{g} = 0 & \text{on } \partial\Omega \end{cases} \tag{4}$$

where $\hat{g} \in L^2(\Omega)$ represents the misfit of the penalized PDE, i.e., $\hat{g} = L\hat{f} - u$, L^* is the adjoint operator of L, i.e., is such that $\int_\Omega L\varphi\psi = \int_\Omega \varphi L^*\psi \ \forall\varphi,\psi \in V$, and is defined as

$$L^*\hat{g} = -div(\mathbf{K}\nabla\hat{g}) - \mathbf{b}\cdot\nabla\hat{g} + (c - div(\mathbf{b}))\hat{g}. \tag{5}$$

The proof of Proposition 1 is based on PDE optimal control theory (see, e.g., [9]) and is detailed in [1], where the regularity conditions required on the parameters of the PDE are also specified; the proof takes into account also the more general boundary conditions described in Section A.2.

3 Model for areal data

We here extend the surface smoothing method presented in the previous Section to the case of areal data, a setting common in many applications, including the one driving our study.

Let $D_i \subset \Omega$, for $i = 1,\dots,N$, be some subdomains where we have observations and z_{ij}, for $j = 1,\dots,n_i$, be the observations located at point $\mathbf{p}_{ij} \in D_i$. For the observations z_{ij}, we consider the pointwise model (1), i.e.,

$$z_{ij} = f_0(\mathbf{p}_{ij}) + \epsilon_{ij} \tag{6}$$

where ϵ_{ij}, for $i = 1,\dots,N$ and $j = 1,\dots,n_i$, are independent errors with zero mean and constant variance σ^2.

In the blood flow velocity application, the location points $\mathbf{p}_{ij}$ are unknown, the only available information being that $\mathbf{p}_{ij} \in D_i$, where D_i is the i-th ECD acquisition beam. We may assume that the location points $\mathbf{p}_{ij}$ are distributed over the subdomains according to a global uniform distribution over Ω and that the subdomains are not overlapping. For each beam D_i, the ECD signal (Figure 1 Left) provides, at a fixed time, a histogram of the measured blood particle velocities. We summarize the information carried by the histogram by its mean value. Specifically, let $\bar{z}_i$ be the mean value of the observations on the subdomain D_i, for $i = 1,\dots,N$. From (6), we can derive the following model for this variable:

$$\bar{z}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} f_0(\mathbf{p}_{ij}) + \frac{1}{n_i}\sum_{j=1}^{n_i}\epsilon_{ij}$$

where $\zeta_i = \sum_{j=1}^{n_i}\epsilon_{ij}/n_i$, $i = 1,\dots,N$, are errors with zero mean and variance σ^2/n_i.

The quantity $\sum_{j=1}^{n_i} f_0(\mathbf{p}_{ij})/n_i$ is the Monte Carlo approximation of $\mathbb{E}[f_0(P)|\ P \in D_i]$ and the latter is in turn equal to the spatial average of the surface on the subdomain D_i, under the assumption of uniformly distributed observation points, i.e.,

$$\frac{1}{n_i}\sum_{j=1}^{n_i} f_0(\mathbf{p}_{ij}) \approx \mathbb{E}\left[\, f_0(P) \middle|\, P \in D_i\right] = \frac{1}{|D_i|}\int_{D_i} f_0.$$

We may thus consider the following model:

$$\bar{z}_i = \frac{1}{|D_i|}\int_{D_i} f_0 + \eta_i \tag{7}$$

where the error terms η_i have zero mean and variances $\bar{\sigma}_i^2$ inversely proportional to the dimension of the beams D_i; this assumption on the variances is coherent with the assumption on location points being distributed on the subdomains according to a uniform distribution (so that in fact the average number of observations on each subdomain is proportional to the dimension of the subdomain). If the subdomains have the same dimension, as it is in fact the case in our application, this simplifies to variances all equal to $\bar{\sigma}^2$.

In order to estimate the surface we hence propose to minimize the penalized sum-of-square-error functional

$$\bar{J}(f) = \sum_{i=1}^{N} \frac{1}{|D_i|}\left(\int_{D_i} (f - \bar{z}_i)\right)^2 + \lambda \int_{\Omega} (Lf - u)^2 \tag{8}$$

with respect to $f \in V$. The first term is now a weighted least-square-error functional for areal data on the subdomains D_i, where the weights are in fact equal to the inverse of the variances $\bar{\sigma}_i^2$, being $\bar{\sigma}_i^2 \propto 1/|D_i|$. Notice that the functional (8) mixes two different kinds of information: the data provide information only on the areal means of the surface f over the subdomains, while the roughness penalty translates the prior knowledge directly on the shape of f.

3.1 Solution to the estimation problem

The estimation problem can be formulated as follows.

Problem 2. *Find $\hat{f} \in V$ such that*

$$\hat{f} = \underset{f \in V}{\operatorname{argmin}}\, \bar{J}(f).$$

Existence and uniqueness of the surface estimator $\hat{f}$ are provided by the following proposition.

Proposition 2. *Under suitable regularity conditions for L, the solution of Problem 2 exists and is unique. The surface estimator $\hat{f}$ is obtained by solving the following system:*

$$\begin{cases} L\hat{f} = u + \hat{g} & \text{in } \Omega \\ \hat{f} = 0 & \text{on } \partial\Omega \end{cases} \qquad \begin{cases} L^*\hat{g} = -\frac{1}{\lambda}\sum_{i=1}^{N} \frac{1}{|D_i|}\mathbb{I}_{D_i}\int_{D_i}(\hat{f} - \bar{z}_i) & \text{in } \Omega \\ \hat{g} = 0 & \text{on } \partial\Omega \end{cases} \tag{9}$$

where $\hat{g} \in L^2(\Omega)$ represents the misfit of the PDE penalized, i.e., $\hat{g} = L\hat{f} - u$, and L^* is the adjoint operator of L.

The proof is similar to the one of Proposition 1 and is detailed in [1].

Remark 1. *All the results presented in this section can be extended to the case of location points distributed on the subdomains according to a general known global distribution μ over Ω, $P \sim \mu$. The quantity $\sum_{j=1}^{n_i} f_0(\mathbf{p}_{ij})/n_i$ is in fact, also in this case, the Monte Carlo approximation of $\mathbb{E}\left[f_0(P) | \, P \in D_i \right]$:*

$$\frac{1}{n_i} \sum_{j=1}^{n_i} f_0(\mathbf{p}_{ij}) \approx \mathbb{E}\left[f_0(P) | \, P \in D_i \right] = \frac{1}{\mu(D_i)} \int_{D_i} f_0(\mathbf{p}) \mu(d\mathbf{p}).$$

Therefore the model for the areal mean on the subdomains becomes:

$$\bar{z}_i = \frac{1}{\mu(D_i)} \int_{D_i} f_0(\mathbf{p}) \mu(d\mathbf{p}) + \eta_i.$$

Under the assumption of non overlapping subdomains, the errors η_i have zero mean and variances inversely proportional to $\mu(D_i)$, which is the probability of sampling a point in the subdomain D_i. The surface estimator $\hat{f}$ can be obtained minimizing the weighted least square functional

$$\bar{J}_\mu(f) = \sum_{i=1}^{N} \frac{1}{\mu(D_i)} \left(\int_{D_i} \left(f(\mathbf{p}) - \bar{z}_i \right) \mu(d\mathbf{p}) \right)^2 + \lambda \int_\Omega (Lf - u)^2$$

with respect to $f \in V$. The weights in the least square term are proportional to the inverse of $\mathbb{V}ar(\bar{z}_i)$, being $\mathbb{V}ar(\bar{z}_i) \propto 1/\mu(D_i)$.

4 Finite Element solution to the estimation problems

The surface estimation problems in the pointwise and areal data frameworks presented respectively in Sections 2 and 3 are infinite dimensional problems and cannot be solved analytically. PDEs are usually solved in a so-called weak sense and, under the regularity conditions required in Propositions 1 and 2, the weak solution is indeed a classical one. This weak problem (or variational problem) is naturally formulated in the space $H_0^1(\Omega)$, which is the space of functions in $L^2(\Omega)$ with first derivatives in $L^2(\Omega)$ and with $f|_{\partial\Omega} = 0$. The weak problem is than usually discretized by means of the Finite Element method, a standard technique used in engineering applications to approximate PDEs (see, e.g., [12]), that provides a basis for piecewise continuous polynomial surfaces over a triangulation of the domain of interest. The discretization of a surface by means of Finite Elements is similar to the discretization of a curve by means of univariate splines, the latter providing a basis for piecewise polynomial curves.

Let $\mathcal{T}_h$ be a triangulation of the domain, where h denotes the characteristic mesh size. Figure 4 Left shows the triangulation considered in the velocity field application. We consider the space V_h^r of piecewise continuous polynomial functions of order $r \geq 1$ over the triangulation:

$$V_h^r = \left\{ v \in C^0(\bar{\Omega}) : v|_{\boldsymbol{\tau}} \in \mathbb{P}^r(\boldsymbol{\tau}) \; \forall \boldsymbol{\tau} \in \mathcal{T}_h \right\}. \tag{10}$$

Let $N_h = \dim(V_h^r)$ and denote by $\psi_1, \ldots, \psi_{N_h}$ the Finite Element basis functions and by $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{N_h}$ the nodes associated to the N_h basis functions. Notice that the

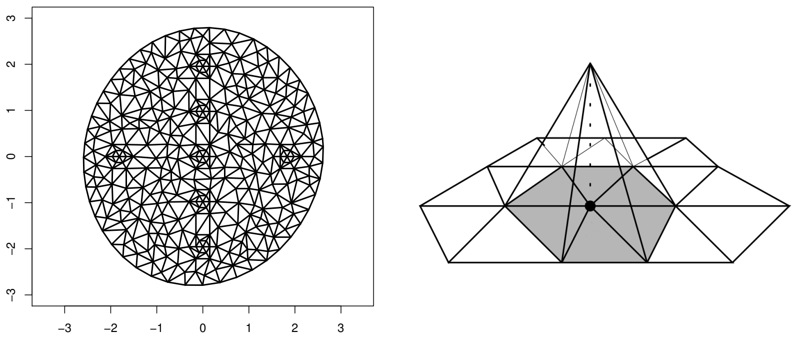

Figure 4: Left: Triangulation of the carotid cross-section of interest in the velocity field application. Right: a linear Finite Element basis function on a triangulation.

mesh can be defined independently of the location points $\mathbf{p}_1, \cdots, \mathbf{p}_n$. The nodes $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_{N_h}$ correspond to the vertices of the triangulation $\mathcal{T}_h$, if the basis is piecewise linear, and are a superset of the vertices when the degree of the polynomial basis is higher than one. Figure 4 Right shows for example a linear Finite Element basis function on a regular triangulation. The basis functions $\psi_1, \ldots, \psi_{N_h}$ are *Lagrangian*, meaning that $\psi_k(\xi_l) = \delta_{\xi_l} \ \forall k = 1, \ldots, N_h$. Hence a surface $f \in V_h^r$ is uniquely determined by its values at the nodes:

$$f(x, y) = \sum_{k=1}^{N_h} f(\xi_k)\psi_k(x, y) = \boldsymbol{\psi}(x, y)^T \mathbf{f}$$

where

$$\mathbf{f} = (f(\boldsymbol{\xi}_1), \ldots, f(\boldsymbol{\xi}_{N_h}))^T$$

and

$$\boldsymbol{\psi} = (\psi_1, \ldots, \psi_{N_h})^T .$$

In the following we consider only homogeneous Dirichlet b.c., for which the value of the function at the boundary is fixed to 0. In this case we consider the Finite Element space $V_{h,0}^r = \left\{ v \in C^0(\bar{\Omega}) : v|_{\partial\Omega} = 0 \text{ and } v|_{\boldsymbol{\tau}} \in \mathbb{P}^r(\boldsymbol{\tau}) \ \ \forall \boldsymbol{\tau} \in \mathcal{T}_h \right\}$, of dimension $N_{h,0}$, which only necessitates of the internal nodes of the triangulation and the associated basis functions, whilst all boundary nodes can be discarded. In Section A.2 we extend the results presented in this section to the case of more general boundary conditions.

4.1 Pointwise estimator

In order to define the weak problem associated to (4) we introduce the bilinear form $a(\cdot, \cdot)$ associated to the operator L, defined as

$$a(\hat{f}, \psi) = \int_\Omega \left(\mathbf{K}\nabla \hat{f} \cdot \nabla \psi + \mathbf{b} \cdot \nabla \hat{f} \psi + c\hat{f}\psi \right). \tag{11}$$

The discrete version of the weak (or variational) problem is thus given by

$$\begin{cases} a(\hat{f}_h, \psi_h) - \int_\Omega \hat{g}_h \psi_h = \int_\Omega u \psi_h \\ \lambda a(\varphi_h, \hat{g}_h) + \sum_{i=1}^n \hat{f}_h(\mathbf{p}_i)\varphi_h(\mathbf{p}_i) = \sum_{i=1}^n z_i \varphi_h(\mathbf{p}_i) \end{cases} \tag{12}$$

for all $\psi_h, \varphi_h \in V_{h,0}^r$, where $\hat{f}_h, \hat{g}_h \in V_{h,0}^r$. This approach allows us to write the estimation problem as a linear system. Define $\boldsymbol{\psi}_x = \left(\partial\psi_1/\partial x, \ldots, \partial\psi_{N_{h,0}}/\partial x\right)^T$ and $\boldsymbol{\psi}_y = (\partial\psi_1/\partial y, \ldots, \partial\psi_{N_{h,0}}/\partial y)^T$ and the matrices

$$\mathbf{R}(c) = \int_\Omega c\boldsymbol{\psi}\boldsymbol{\psi}^T, \quad \mathbf{R}_x(\mathbf{b}) = \int_\Omega \mathbf{b}_1\boldsymbol{\psi}\boldsymbol{\psi}_x^T, \quad \mathbf{R}_y(\mathbf{b}) = \int_\Omega \mathbf{b}_2\boldsymbol{\psi}\boldsymbol{\psi}_y^T, \tag{13}$$

$$\mathbf{R}_{xx}(\mathbf{K}) = \int_\Omega \mathbf{K}_{11}\boldsymbol{\psi}_x\boldsymbol{\psi}_x^T, \quad \mathbf{R}_{yy}(\mathbf{K}) = \int_\Omega \mathbf{K}_{22}\boldsymbol{\psi}_y\boldsymbol{\psi}_y^T, \tag{14}$$

$$\mathbf{R}_{xy}(\mathbf{K}) = \int_\Omega \mathbf{K}_{12}(\boldsymbol{\psi}_x\boldsymbol{\psi}_y^T + \boldsymbol{\psi}_y\boldsymbol{\psi}_x^T), \tag{15}$$

where $\mathbf{K}_{ij}$ and $\mathbf{b}_j$ are the elements of the diffusion tensor matrix $\mathbf{K}$ and of the transport vector $\mathbf{b}$. Using this notation, the Finite Element matrix associated to the bilinear form $a(\cdot,\cdot)$ in (11) is given by

$$\mathbf{A}(\mathbf{K}, \mathbf{b}, c) = \mathbf{R}_{xx}(\mathbf{K}) + \mathbf{R}_{xy}(\mathbf{K}) + \mathbf{R}_{yy}(\mathbf{K}) + \mathbf{R}_x(\mathbf{b}) + \mathbf{R}_y(\mathbf{b}) + \mathbf{R}(c). \tag{16}$$

Moreover, define the vectors $\mathbf{z} = (z_1, \ldots, z_n)^T$, $\mathbf{u} = \int_\Omega u\boldsymbol{\psi}$ and the matrices

$$\mathbf{R} = \mathbf{R}(1) = \int_\Omega \boldsymbol{\psi}\boldsymbol{\psi}^T$$

and

$$\boldsymbol{\Psi} = \begin{bmatrix} \boldsymbol{\psi}^T(\mathbf{p}_1) \\ \vdots \\ \boldsymbol{\psi}^T(\mathbf{p}_n) \end{bmatrix} \tag{17}$$

where $\boldsymbol{\Psi}$ is the matrix of basis evaluations at the n data locations $\mathbf{p}_1, \cdots, \mathbf{p}_n$. The discrete surface estimator is thus provided by the following Proposition.

Proposition 3. *The Finite Element solution $\hat{f}_h$ of the discrete counterpart (12) of the estimation Problem 1 exists, is unique and is given by $\hat{f}_h = \boldsymbol{\psi}^T\hat{\mathbf{f}}$ where $\hat{\mathbf{f}}$ is the solution of the linear system*

$$\begin{bmatrix} \boldsymbol{\Psi}^T\boldsymbol{\Psi} & \lambda\mathbf{A}^T \\ \mathbf{A} & -\mathbf{R} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Psi}^T\mathbf{z} \\ \mathbf{u} \end{bmatrix}. \tag{18}$$

The proof of well-posedness of the discrete problem is given in [1].

4.1.1 Properties of the estimator

The estimator $\hat{f}_h$ is a linear function of the observed data values. The fitted values $\hat{\mathbf{z}} = \boldsymbol{\Psi}\hat{\mathbf{f}}$ can be represented as

$$\hat{\mathbf{z}} = \mathbf{S}\mathbf{z} + \mathbf{r} \tag{19}$$

where the smoothing matrix $\mathbf{S} \in \mathbb{R}^{n\times n}$ and the vector $\mathbf{r} \in \mathbb{R}^n$ are obtained as

$$\mathbf{S} = \boldsymbol{\Psi}\left(\boldsymbol{\Psi}^T\boldsymbol{\Psi} + \lambda\mathbf{P}\right)^{-1}\boldsymbol{\Psi}^T, \tag{20}$$

$$\mathbf{r} = \boldsymbol{\Psi}\left(\boldsymbol{\Psi}^T\boldsymbol{\Psi} + \lambda\mathbf{P}\right)^{-1}\lambda\mathbf{P}\mathbf{A}^{-1}\mathbf{u}. \tag{21}$$

with $\mathbf{P}$ denoting the penalty matrix

$$\mathbf{P} = \mathbf{P}(\mathbf{K}, \mathbf{b}, c) = \mathbf{A}^T (\mathbf{R})^{-1} \mathbf{A}. \tag{22}$$

The smoothing matrix $\mathbf{S}$ has the typical form obtained in a penalized regression problem. In particular, the positive definite penalty matrix $\mathbf{P}$ represents the discretization of the penalty term in (2). Notice that, thanks to the weak formulation of the estimation problem, this penalty matrix does not involve the computation of second order derivatives. Appendix A.1 shows that, on the Finite Element space used to discretize the problem, $\mathbf{P}$ is in fact analogue to the penalty matrix $\tilde{\mathbf{P}}$ that would be obtained as direct discretization of the penalty term in (2), involving the computation of second order derivatives. Finally, the vector $\mathbf{r}$ is equal to zero when the penalized PDE is homogeneous ($u = 0$); notice that when no specific information on the forcing term is available, it is indeed preferable to consider homogeneous PDEs.

Thanks to the linearity of the estimator $\hat{\mathbf{z}}$ in the observations we can easily derive its properties and obtain classical inferential tools as pointwise confidence bands and prediction intervals (see also [16]). Let $\mathbf{z}_0 = (f_0(\mathbf{p}_1), \ldots, f_0(\mathbf{p}_n))^T$ be the column vector of evaluations of the true function f_0 at the n data locations. Recalling that $\mathbb{E}[\mathbf{z}] = \mathbf{z}_0$ and $\mathbb{C}\mathrm{ov}(\mathbf{z}) = \sigma^2 \mathbf{I}$ in our model definition, we can compute the expected value and the variance of the estimator $\hat{\mathbf{z}}$:

$$\mathbb{E}[\hat{\mathbf{z}}] = \mathbf{S}\mathbf{f}_0 + \mathbf{b} \qquad \text{and} \qquad \mathbb{C}\mathrm{ov}(\hat{\mathbf{z}}) = \sigma^2 \mathbf{S}\mathbf{S}^T.$$

Since we are dealing with linear estimators, we can use $tr(\mathbf{S})$ as a measure of the equivalent degrees of freedom for linear estimators (see, e.g., [2] and [7]). Hence we can estimate σ^2 as

$$\hat{\sigma}^2 = \frac{1}{n - tr(\mathbf{S})} (\hat{\mathbf{z}} - \mathbf{z})^T (\hat{\mathbf{z}} - \mathbf{z}).$$

The smoothing parameter λ may be selected via Generalized Cross-Validation minimizing the index

$$GCV(\lambda) = \frac{1}{n\,(1 - tr(\mathbf{S})/n)^2} (\hat{\mathbf{z}} - \mathbf{z})^T (\hat{\mathbf{z}} - \mathbf{z}).$$

4.2 Areal estimator

Analogously to the case of pointwise observations, also with areal observations we can introduce an equivalent variational formulation of the estimation problem. Specifically, the variational problem associated to (9) can be discretized as

$$\begin{cases} a(\hat{f}_h, \psi_h) - \int_\Omega \hat{g}_h \psi_h = \int_\Omega u \psi_h \\ \lambda a(\varphi_h, \hat{g}_h) + \sum_{i=1}^N \frac{1}{|D_i|} \int_{D_i} \hat{f}_h \int_{D_i} \varphi_h = \sum_{i=1}^N \bar{z}_i \int_{D_i} \varphi_h \end{cases} \tag{23}$$

for all $\psi_h, \varphi_h \in V_{h,0}^r$, where $\hat{f}_h, \hat{g}_h \in V_{h,0}^r$ and $a(\cdot,\cdot)$ is the bilinear form defined in (11).

Let $\bar{\mathbf{z}} = (\bar{z}_1, \ldots, \bar{z}_N)^T$ be the vector of mean values on subdomains $D_i, \ldots, D_N$, and

$$\bar{\mathbf{\Psi}} = \begin{bmatrix} \frac{1}{|D_1|} \int_{D_1} \boldsymbol{\psi}^T \\ \vdots \\ \frac{1}{|D_N|} \int_{D_N} \boldsymbol{\psi}^T \end{bmatrix}$$

be the matrix of spatial means of the basis functions on the subdomains; moreover, introduce the weight matrix $\mathbf{W} = \mathrm{diag}(|D_1|, \ldots, |D_N|)$ (recall that $\bar{\sigma}_i^2 \propto 1/|D_i|$). The existence and the uniqueness of the discrete surface estimator is stated by the following Proposition.

Proposition 4. *The Finite Element solution $\hat{f}_h$ of the discrete counterpart of the estimation Problem 2 exists, is unique and is given by $\hat{f}_h = \boldsymbol{\psi}^T \hat{\mathbf{f}}$ where $\hat{\mathbf{f}}$ is the solution of the linear system*

$$\begin{bmatrix} \bar{\mathbf{\Psi}}^T \mathbf{W} \bar{\mathbf{\Psi}} & \lambda \mathbf{A^T} \\ \mathbf{A} & -\mathbf{R} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{\Psi}}^T \mathbf{W} \bar{\mathbf{z}} \\ \mathbf{u} \end{bmatrix}. \tag{24}$$

The proof of well-posedness of the discrete problem is given in [1].

Notice that even if the method provides a pointwise surface estimator $\hat{f}_h$, in the areal data framework we are instead interested in the estimator of the spatial mean of the surface on a subdomain D:

$$\bar{\hat{f}}\,(D) = \frac{1}{|D|} \int_D \hat{f}$$

The Finite Element counterpart of this estimator is defined as

$$\bar{\hat{f}}_h\,(D) = \frac{1}{|D|} \int_D \hat{f}_h = \bar{\boldsymbol{\psi}}_D^T \hat{\mathbf{f}}$$

where $\bar{\boldsymbol{\psi}}_D = (1/|D| \int_D \psi_1, \ldots, 1/|D| \int_D \psi_{N_{h,0}})^T$.

4.2.1 Properties of the estimator

The discrete surface estimator $\hat{f}_h$ and the estimator of the spatial average on the subdomains $\bar{\hat{f}}_h$ are linear in the observed data values $\bar{\mathbf{z}}$. The fitted values of the spatial average on the subdmains $D_1, \ldots, D_N$ are defined as $\hat{\bar{\mathbf{z}}} = \bar{\mathbf{\Psi}} \hat{\mathbf{f}} = (\bar{\hat{f}}_h\,(D_1), \ldots, \bar{\hat{f}}_h\,(D_N))^T$. They can be represented as

$$\hat{\bar{\mathbf{z}}} = \bar{\mathbf{S}} \bar{\mathbf{z}} + \bar{\mathbf{r}} \tag{25}$$

where $\bar{\mathbf{S}} \in \mathbb{R}^{N \times N}$ and $\bar{\mathbf{r}} \in \mathbb{R}^N$ are defined as

$$\bar{\mathbf{S}} = \bar{\mathbf{\Psi}} \left(\bar{\mathbf{\Psi}}^T \mathbf{W} \bar{\mathbf{\Psi}} + \lambda \mathbf{P} \right)^{-1} \bar{\mathbf{\Psi}}^T \mathbf{W}, \tag{26}$$

$$\mathbf{b} = \bar{\mathbf{\Psi}} \left(\bar{\mathbf{\Psi}}^T \mathbf{W} \bar{\mathbf{\Psi}} + \lambda \mathbf{P} \right)^{-1} \lambda \mathbf{P} \mathbf{A}^{-1} \mathbf{u}. \tag{27}$$

From the definition of model (7) and the linearity of the estimator we can derive the mean of the estimator

$$\mathbb{E}\left[\hat{\bar{\mathbf{z}}}\right] = \bar{\mathbf{S}} \bar{\mathbf{z}}_0 + \bar{\mathbf{b}}, \tag{28}$$

where $[\bar{\mathbf{z}}_0]_i = 1/|D_i| \int_{D_i} f_0$, and its covariance

$$\mathbb{C}\mathrm{ov}(\hat{\bar{\mathbf{z}}}) = \bar{\mathbf{S}}\ \mathrm{diag}(\bar{\sigma}_1^2, \ldots, \bar{\sigma}_N^2)\ \bar{\mathbf{S}}^T. \tag{29}$$

It should be noticed that in the areal data framework the expected value (28) and the variance (29) refer to the estimator of the spatial mean on a subdomain. In fact, even though we can obtain a pointwise estimator for the surface $\hat{f}_h$ as described in Proposition 4, we cannot provide an accurate uncertainty quantification for this estimate, because model (7) provides information only on the areal errors η_i. In particular, in the considered areal framework, the variance

$$\mathbb{C}\text{ov}(\hat{\mathbf{f}}) = \mathbf{\Psi}\bar{\mathbf{\Psi}}^{-1}\bar{\mathbf{S}}\ \text{diag}(\bar{\sigma}_1^2, \dots, \bar{\sigma}_N^2)\ \bar{\mathbf{S}}^T\bar{\mathbf{\Psi}}^{-T}\mathbf{\Psi}^T$$

would underestimate the real variance of $\hat{\mathbf{f}}$.

5 General boundary conditions

All the results presented in Sections 2, 3 and 4 can be extended to the case of general homogeneous and non-homogeneous boundary conditions involving the value of the surface or of its first derivatives at the boundary $\partial\Omega$, allowing for a complex modeling of the phenomenon behavior at the boundary of the domain. The three classic boundary conditions for second order PDEs are Dirichlet, Neumann and Robin conditions. The Dirichlet condition controls the value of the function at the boundary, i.e., $f|_{\partial\Omega} = h_D$, the Neumann condition concerns the value of the normal derivative of the function at the boundary, i.e., $\mathbf{K}\nabla f \cdot \boldsymbol{\nu}|_{\partial\Omega} = h_N$, where $\boldsymbol{\nu}$ is the outward unit normal vector to $\partial\Omega$, while the Robin condition involves the value of a linear combination of first derivative and the value of the function at the boundary, i.e., $\mathbf{K}\nabla f \cdot \boldsymbol{\nu} + \gamma f|_{\partial\Omega} = h_R$. We can also impose different boundary conditions on different portions of the boundary that form a partition of $\partial\Omega$. All the admissible boundary conditions can be summarized as

$$\begin{cases} f = h_D & \text{on } \Gamma_D \\ \mathbf{K}\nabla f \cdot \boldsymbol{\nu} = h_N & \text{on } \Gamma_N \\ \mathbf{K}\nabla f \cdot \boldsymbol{\nu} + \gamma f = h_R & \text{on } \Gamma_R \end{cases} \tag{30}$$

where h_D, h_N and h_R have to satisfy some regularity conditions in order to obtain a well defined functional $J(f)$ (see, e.g., [4]).
Under (30), the solution of the estimation problem and of its discrete counterpart also involve boundary terms. Appendix A.2 gives all the details for this general case.

6 Simulation studies

In this Section we study the performances of the SR-PDE, comparing it to standard SSR and to SOAP in simple simulation studies that mimic our application setting. The domain Ω is quasi circular; the true surface f_0, represented in Figure 5, is obtained as a deformation of a parabolic profile using landmark registration and is equal to zero at the boundary of the domain. Likewise for our application, we assume to have a priori information about the shape of the field, that is known to have a quasi parabolic profile, with almost circular isolines, and to be zero at the boundary.

Since SOAP is not currently devised to deal with areal data, we consider here pointwise observations, with location points sampled on the whole or only on subregions of the domain. Specifically, we consider three cases:

A. n=100 observation points $\mathbf{p}_1, \dots, \mathbf{p}_n$ uniformly sampled on the entire domain;

B. n=100 observation points uniformly sampled only on the first and third quadrants;

C. n=100 observation points sampled in a cross-shape pattern.

The experiment is replicated 50 times. For each study case, A, B, and C, and each replicate: we sample the location points, $\mathbf{p}_1, \ldots, \mathbf{p}_n$; we sample independent errors, $\epsilon_1, \ldots, \epsilon_n$, from a Gaussian distribution with mean 0 and standard deviation $\sigma = 0.1$; we thus obtain observations $z_1, \ldots, z_n$ from model (1) with the true function f_0 displayed in Figure 5.

The surface $\hat{f}$ is estimated using three methods:

1. SR-PDE smoothing (anisotropic smoothing);
2. standard SSR (isotropic smoothing);
3. SOAP (isotropic smoothing).

For all the three methods, we impose homogeneous Dirichlet b.c., $f|_{\partial\Omega} = 0$; for each simulation study, each replicate and each method, the value of the smoothing parameter λ is chosen via GCV.

The triangulation used for the SR-PDE and standard SSR estimation is a uniform mesh on the domain, represented in Figure 5 Right, with approximately 100 vertices. Both for SR-PDE and SSR we use a linear Finite Element space for the discretization of the surface estimator.

Using SR-PDE it is possible to incorporate the prior knowledge on the shape of the surface, that should have almost circular isolines. We can achieve this by penalizing a PDE that smooths the surface along concentric circles; specifically we consider the anisotropic diffusion tensor

$$\mathbf{K}(x, y) = \begin{bmatrix} y^2 + \kappa_1 x^2 & (\kappa_1 - 1)xy \\ (\kappa_1 - 1)xy & x^2 + \kappa_1 y^2 \end{bmatrix} + \kappa_2 \left(R^2 - x^2 - y^2\right) \mathbf{I}_2, \tag{31}$$

where R denotes the largest radius in this almost circular domain (in these simulations, $R = 1$) and we set $\kappa_1 = 0.01$, $\kappa_2 = 0.1$; this diffusion tensor is shown in the right panel of Figure 5. The first hyperparameter represents the ratio between the diffusion in the radial and in the circular direction. The anisotropic part of the diffusion field, which corresponds to the first term of the right-hand side of (31),

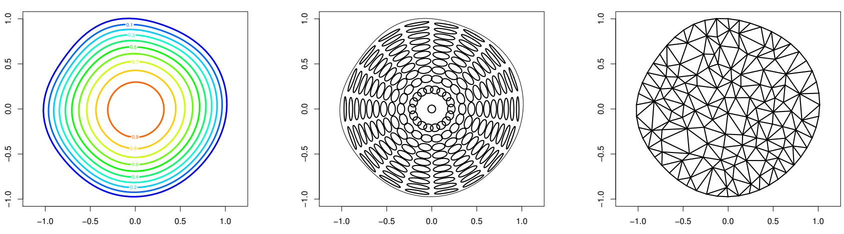

Figure 5: Left: true surface f_0, with almost circular isolines and zero value at the boundary of the domain, used for the simulation studies; the image displays the isolines $(0, 0.1, \ldots, 0.9, 1)$. Center: diffusion tensor field $\mathbf{K}$ used in SR-PDE. Right: triangulation of the domain Ω used in SSR and SR-PDE.

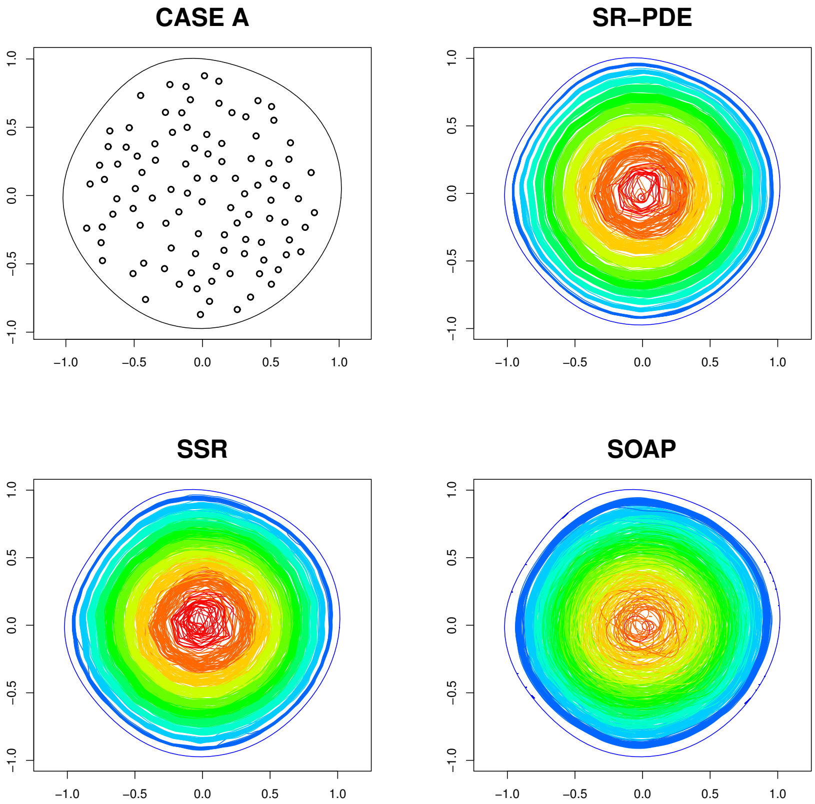

Figure 6: Top left: location points sampled in the first replicate for case A. Top right, bottom left, bottom right: surface estimates obtained using respectively SR-PDE, SSR and SOAP; the images display the isolines $(0, 0.1, \ldots, 0.9, 1)$ of the surface estimates obtained in the 50 simulation replicates; the isolines are colored using the same color scale used for the isolines of the true function f_0 in Figure 5.

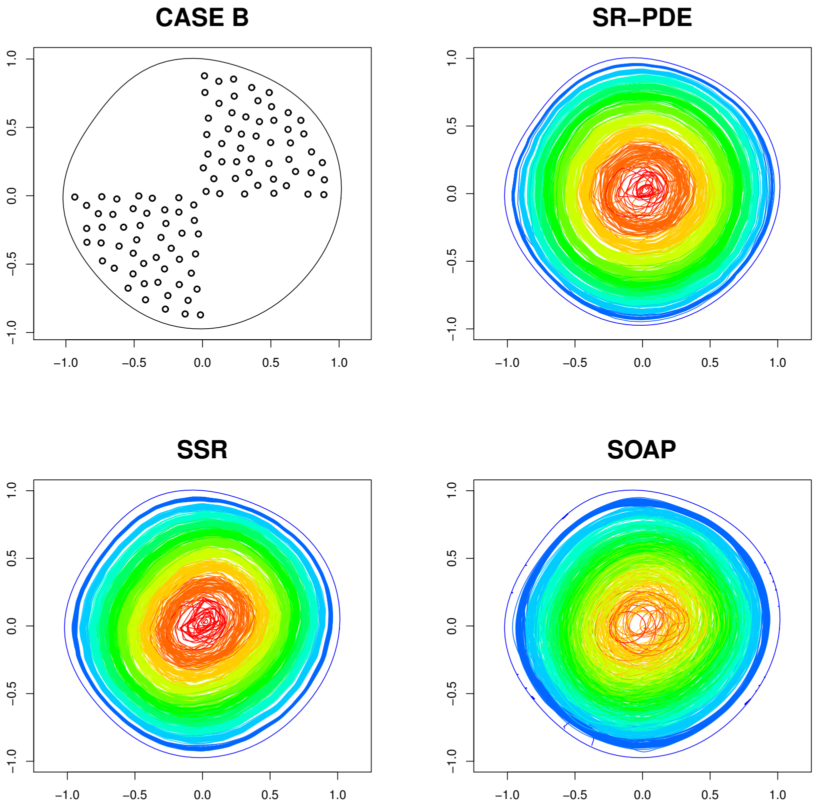

Figure 7: Same as Figure 6, for case B.

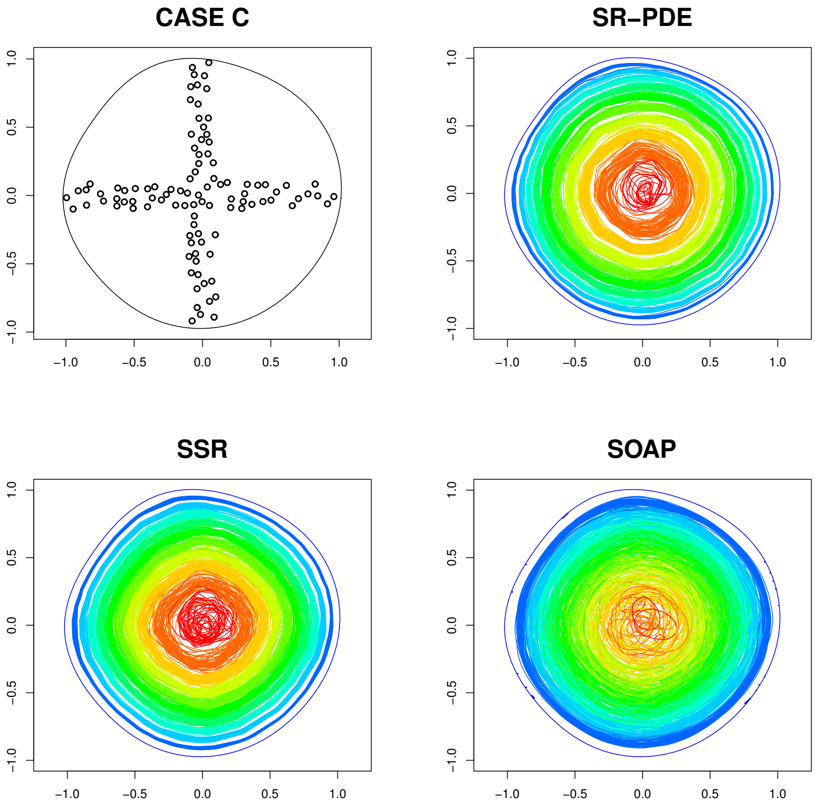

Figure 8: Same as Figure 6, for case C.

is stronger near the boundary and completely vanishes in the center of the carotid; instead the isotropic part, which corresponds to the second term of the right-hand side of (31), vanishes near the boundary. The relative strength of the anisotropic and isotropic part is controlled via κ_2. The transport field, the reaction term and the forcing term are set equal to zero, i.e., $\mathbf{b} = 0$, $c = 0$ and $u = 0$.

Standard SSR instead is not able to take advantage of the specific prior knowledge of the shape of the surface, and enforces an isotropic smoothing, corresponding to SR-PDE with $\mathbf{K} = \mathbf{I}$, $\mathbf{b} = 0$, $c = 0$ and $u = 0$. Also SOAP produces an isotropic smoothing; this technique is implemented using the function `gam`, in the R package `mgcv` 1.7-22, see [18], using 49 interior knots on a lattice.

Figures 6-8 show the results obtained using the different methods in the three considered scenarios, cases A, B and C. The upper left panel of the figures shows location points sampled in the first replicate in each of the three different scenarios. The top right, bottom left, bottom right panels of these Figures display the surface estimates obtained using respectively SR-PDE, SSR and SOAP. In particular, the images display the isolines $(0, 0.1, \ldots, 0.9, 1)$ of the surface estimates obtained in the 50 simulation replicates; the isolines are colored using the same color scale used for the isolines of the true function f_0 in Figure 5.

Comparing the results obtained with the three methods we can notice that the inclusion of the prior knowledge improves the estimate, especially when data are distributed only on subregions of the domain. We can in fact see that in the three case studies the surfaces estimated with SR-PDE smoothing have circular isolines similar to those of the true surface f_0. Instead, when the prior knowledge is not included in the model, i.e., for standard SSR and SOAP, the surface estimates tend to depend on the design of the experiments. We notice in fact that the isolines of SSR and SOAP estimates are similar to ellipses in case B and to rhomboids in case C, instead of circles. This is due to the fact that both methods tend to fit planes in those areas where no observations are available. This phenomenon is more apparent with SSR than with SOAP because SOAP estimates have an higher variability.

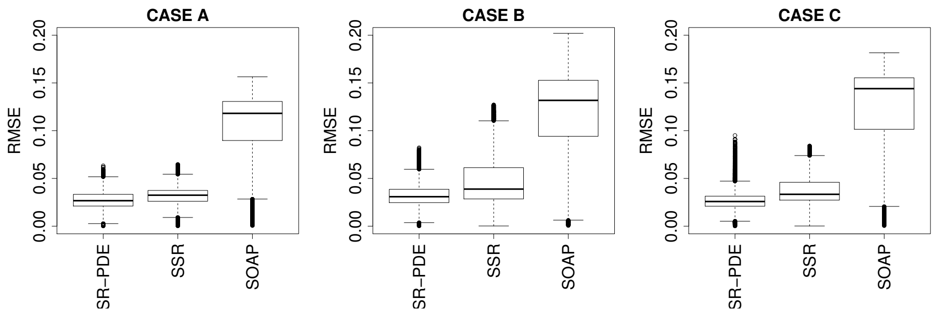

Figure 9: Boxplot of RMSE (evaluated on a fine lattice of step 0.01 over the domain Ω) for SR-PDE, SSR and SOAP estimators, in case studies A, B and C (left, central and right panel, respectively).

Figure 9 shows the comparison of the three methods in terms of root mean square error (RMSE) of the corresponding estimators, with the RMSE evaluated on a fine lattice of step 0.01 over the domain Ω. The boxplots highlight that incorporation of the prior knowledge on the shape of the surface leads to a large improvement in the estimation. SR-PDE smoothing provides in fact significantly better estimates of f_0 than the other two methods. The boxplots also show that SR-PDE estimates display lower variability than SSR and SOAP estimates. This phenomenon is also visible from the isolines of the estimated surfaces with SR-PDE, SSR and SOAP represented Figures 6-8.

7 Application to the blood-flow velocity field estimation

Carotid ECD is usually the first imaging procedure used to diagnose carotid artery diseases, such as ischemic stroke, caused by the presence of an atherosclerotic plaque. ECD data in our study have been collected using a Diagnostic Ultrasound System Philips iU22 (Philips Ultrasound, Bothell, U.S.A.) with a L12-5 probe. The septum that divides the carotid bifurcation is localized and marked as a reference point. With the help of an electronic rule, we localize the other points of acquisition of the blood velocity; specifically, in our protocol the blood flow velocity is measured in standard locations points, according to the cross-shaped design represented in the right panel of Figure 1, on the carotid cross-section located 2 cm before the reference point indicated above.

In order to estimate the systolic velocity field on this cross-section of the carotid we minimize the functional $\bar{J}(f)$ defined in (8). As mentioned in Section 1 we know that a physiological velocity profile has smooth and almost circular isolines. For this reason we choose to penalize a PDE that includes the space varying anisotropic diffusion tensor shown in the left panel of Figure 3 and described in equation (31) (where the largest section radius is $R = 2.8$ and we set $\kappa_1 = 0.1$, $\kappa_2 = 0.2$), that smooths the observations in the tangential direction of concentric circles. Moreover, we also know that, due to viscosity of the blood, a physiological velocity field is rather flat on the central part of the artery lumen. For this reason, we also include in the PDE model the space varying transport field shown in the right panel of Figure 3, which smooths the observations in the radial direction, from the center of the cross-section to the boundary: $\mathbf{b}(x, y) = (\beta x, \beta y)^T$, where the hyperparameter β represents the intensity of the transport field (here we set $\beta = 0.5$). This transport term in fact penalizes high first derivatives in the radial direction, providing velocity profiles that tend to flatten in the central part of the artery lumen. The reaction parameter and the forcing term are not needed in this application, hence we set $c = 0$ and $u = 0$. Finally, we know that blood flow velocity is zero at the arterial wall, due to friction between the blood particles and the vessel wall (the above mentioned no-slip conditions) and hence we impose homogeneous Dirichlet b.c.: $f|_{\partial\Omega} = 0$. The problem is then discretized by means of linear Finite Elements defined on the mesh represented in the left panel of Figure 4.

Figure 10 displays the velocity field estimated using SR-PDE smoothing. A visual comparison with the estimate obtained for the same data by standard SSR, shown in Figure 2, immediately highlights the advantages of the proposed technique. Whilst the standard SSR estimate is strongly influenced by the cross-shaped pattern of the observations and displays strongly rhomboidal isolines, forcing the surface estimate towards a plane in regions where no observations are available, the SR-PDE efficiently uses the a priori information on the phenomenon under study and returns a realistic estimate of the blood flow, which is not affected by the cross-shaped pattern of the observations and displays physiological almost circular isolines.

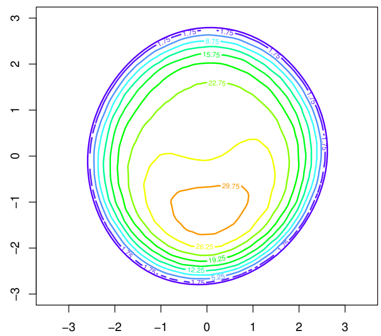

Figure 10: Estimate of the blood-flow velocity field in the carotid section with SR-PDE.

Notice that the SR-PDE estimate captures an asymmetry in the data, resulting in an eccentric estimate of the blood flow: the velocity peak is in fact not in the center of the cross-section but in the lower part where higher velocities are measured. This feature of the blood flow is indeed justified by the curvature of the carotid artery and by the non-stationarity of the blood flow. SR-PDE estimates in fact accurately highlight important features of the blood flow, such as eccentricity, asymmetry and reversion of the fluxes, that are of interest to the medical doctors, in order to understand how the local hemodynamics influences atherosclerosis pathogenesis. As mentioned in the Introduction, MACAREN@MOX project aims in fact at exploring this relationship, investigating how different hemodynamical patterns affect the plaque formation process. For this reason, obtaining accurate physiological estimates of blood flow velocity fields is a first crucial goal of the project. Indeed, the SR-PDE estimates will then be used in populations studies that compare the blood flow velocity field in patients vs healthy subjects, and that compare the velocity field in patients before and after the removal of the carotid plaque via thromboendarterectomy. Notice that such population studies involve the comparisons of estimates referred to different domains, since the cross-sections of the carotids have of course patient-specific shapes; to face this issue we are currently developing an appropriate registration method (see Section 8) and these analysis will be the object of a following dedicated work. The estimated velocity fields will also be used as inflow conditions for the hemodynamic simulations performed using the patient-specific carotid morphology. The prescription of suitable inflow conditions in computational fluid-dynamics is in fact a major issue; see, e.g., [17]. Moreover, the computation of the variance of the surface estimator will also be used to investigate the sensitivity of these simulations to the specified inflow conditions and will provide some understanding on how their misspecification affects the results. These numerical simulations will in turn offer enhanced data that give a richer information on hemodynamical regimes in the carotid bifurcation, further allowing the study of its impact on atherosclerosis. Computational fluid-dynamic simulations are also of great interest because they allow to synthetically verify the impact of different surgical interventions, evaluating which one is more prone to the reformation of the plaque or to other complications. In the future, this could become an important tool for comparing beforehand the effects of different interventions for a given patient, with respect to the geometry of the patient carotid and to the properties of the atherosclerotic plaque, giving important suggestions to clinicians on the surgical operation to choose in different situations.

8 Conclusion and future work

In this work we have introduced an innovative method for surface and spatial field estimation, when prior knowledge is available, concerning the physics of the problem. In particular, this prior knowledge, conveniently described via a PDE, is used to model the space variation of the phenomenon. Although demonstrated on the specific application that motivated its development, the method has indeed a very broad applicability, since PDEs are commonly used to model phenomena behavior in many fields of sciences and engineering.

One of the most interesting developments within this line of research consists now in the data driven estimation of the hyperparameters in the penalized PDE. In the current study, these hyperparameters have in fact been considered fixed. Notice that, while a currently crucial topic in statistics concerns the development of methods for parameter estimations in Ordinary Differential Equation, this would instead consist in approaching the remarkably more complex problem of data driven estimation of the parameters in PDEs, a research field still largely unexplored by statisticians. To face such problem a possible road is offered by the parameter cascading methodology proposed in [14].

As derived in Section 4, the proposed estimators are linear in the observed data values and have a typical penalized regression form, so that important distributional properties can be readily derived. We are currently also studying the (infill) asymptotic properties of these estimators, when the number of observations n goes to infinity and the characteristic mesh size h goes to zero. Convergence of the estimator when h goes to zero is detailed in [1].

The proposed method can also be extended to include the time dimension, in order to model surfaces evolving in time. Such extension would allow to study how the blood-flow velocity field varies during the time of the heart beat. Notice that it is necessary in this case to allow for changes of the shape of the domain over time, to account for the deformation of the artery wall during the heart beat. This poses a problem of registration of different domains similar to the one faced in population studies (see Section 7).

Finally, the method could also be extended to Riemannian manifold domains, by appropriately setting the problem in the framework presented in [3].

Acknowledgments We would like to thank Christian Vergara, principal investigator of the MACAREN@MOX project, Elena Faggiano for the reconstruction of carotid geometry from MRI data, and Silvia Romagnoli for the acquisition of ecodoppler data. This work is supported by MIUR Ministero dell'Istruzione dell'Università e della Ricerca, *FIRB Futuro in Ricerca* research project "Advanced statistical and numerical methods for the analysis of high dimensional functional data in life sciences and engineering" (http://mox.polimi.it/users/sangalli/firbSNAPLE.html) and by the program Dote Ricercatore Politecnico di Milano - Regione Lombardia, research project "Functional data analysis for life sciences".

A Appendix

A.1 The penalization matrix

Assume for the moment that ψ_j are smooth functions and neglect the forcing term u. The penalty matrix $\mathbf{P}$ in (22) can be written as

$$\mathbf{P} = \int_\Omega \int_\Omega L\boldsymbol{\psi}(\mathbf{s})\boldsymbol{\psi}^T(\mathbf{s}) \left[\int_\Omega \boldsymbol{\psi}\boldsymbol{\psi}^T\right]^{-1} \boldsymbol{\psi}(\mathbf{t})L\boldsymbol{\psi}^T(\mathbf{t})d\mathbf{s}d\mathbf{t}$$

where $L\boldsymbol{\psi} = \left(L\psi_1, \ldots, L\psi_{N_{h,0}}\right)^T$, since $\mathbf{A}_{ij} = a(\psi_j, \psi_i) = \int_\Omega \psi_i L\psi_j$. The matrix $\tilde{\mathbf{P}} = \int_\Omega L\boldsymbol{\psi}L\boldsymbol{\psi}^T$, which may instead be obtained as direct discretization of the penalty term in (2), can be represented as

$$\tilde{\mathbf{P}} = \int_\Omega \int_\Omega L\boldsymbol{\psi}(\mathbf{s})\delta(\mathbf{s},\mathbf{t})L\boldsymbol{\psi}^T(\mathbf{t})d\mathbf{s}d\mathbf{t}$$

using the kernel operator associated with the L^2 space, $\delta(\mathbf{s},\mathbf{t})$, defined as

$$\int_\Omega \delta(\mathbf{s},\mathbf{t})q(\mathbf{t})d\mathbf{t} = q(\mathbf{s}) \quad \forall q \in L^2(\Omega) \cap C(\Omega). \tag{32}$$

From the above equations we see that $\mathbf{P}$ is an approximation in a weak sense of $\tilde{\mathbf{P}}$. In fact, the operator $\delta(\mathbf{s},\mathbf{t})$ is approximated, in the mixed Finite Element approach here considered, with the projection operator

$$\boldsymbol{\psi}^T(\mathbf{s}) \left[\int_\Omega \boldsymbol{\psi}\boldsymbol{\psi}^T\right]^{-1} \boldsymbol{\psi}(\mathbf{t})$$

that projects functions on span$\{\psi_1, \ldots, \psi_{N_{h,0}}\}$. This operator satisfies the property (32) in span$\{\psi_1, \ldots, \psi_{N_h}\}$; in fact, if $q(\mathbf{t}) = \sum_{k=1}^K q_k\psi_k(\mathbf{t})$, then

$$\int_\Omega \boldsymbol{\psi}^T(\mathbf{s}) \left[\int_\Omega \boldsymbol{\psi}\boldsymbol{\psi}^T\right]^{-1} \boldsymbol{\psi}(\mathbf{t})q(\mathbf{t})d\mathbf{t} = \sum_{k=1}^K q_k\psi_k(\mathbf{s}) = q(\mathbf{s})$$

while if $q \notin$ span$\{\psi_1, \ldots, \psi_{N_{h,0}}\}$ this operator projects the function q on span$\{\psi_1, \ldots, \psi_{N_{h,0}}\}$.

A.2 General boundary conditions

In the case of general boundary conditions, the space V is the space of functions in $L^2(\Omega)$ with first and second derivatives in $L^2(\Omega)$ that satisfy (30).

Starting with the pointwise data framework, the estimation problem with general b.c. (30) is analogous to Problem 1. The unique solution of the problem, $\hat{f} \in V$, is obtained by solving

$$\begin{cases} L\hat{f} = u + \hat{g} & \text{in } \Omega \\ + \, b.c. & \text{on } \partial\Omega \end{cases} \qquad \begin{cases} L^*\hat{g} = -\frac{1}{\lambda}\sum_{i=1}^n(\hat{f} - z_i)\delta_{\mathbf{p}_i} & \text{in } \Omega \\ + \, b.c.^* & \text{on } \partial\Omega \end{cases} \tag{33}$$

where $\hat{g} \in L^2(\Omega)$ represents the misfit of the penalized PDE, L^* is the adjoint operator of L and b.c.* are the boundary conditions associated to the adjoint problem, i.e.,

$$\begin{cases} g=0 & \text{on } \Gamma_D \\ \mathbf{K}\nabla g\cdot\boldsymbol{\nu}+\mathbf{b}\cdot\boldsymbol{\nu} g=0 & \text{on } \Gamma_N \\ \mathbf{K}\nabla g\cdot\boldsymbol{\nu}+(\mathbf{b}\cdot\boldsymbol{\nu}+\gamma)g=0 & \text{on } \Gamma_R. \end{cases} \tag{34}$$

Notice that these conditions are always homogeneous.

We define now the space V^r_{h,Γ_D} of piecewise continuous polynomial functions of degree $r\geq 1$ on the domain triangulation, that vanish on Γ_D, the part of the boundary $\partial\Omega$ with Dirichlet b.c. (if Γ_D is not empty):

$$V^r_{h,\Gamma_D}=\left\{v\in C^0(\bar{\Omega}): v|_{\Gamma_D}=0 \text{ and } v|_{\boldsymbol{\tau}}\in\mathbb{P}^r(\boldsymbol{\tau})\ \forall\boldsymbol{\tau}\in\mathcal{T}_h\right\}.$$

We denote by $\psi_1,\ldots,\psi_{N_{h,\Gamma_D}}$, where $N_{h,\Gamma_D}=\dim(V^r_{h,\Gamma_D})$, the Finite Element basis functions of this space, and by $\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_{N_{h,\Gamma_D}}$ the associated nodes; note that the nodes now include the internal nodes and the nodes on Γ_N and Γ_R. Correspondingly the basis vector $\boldsymbol{\psi}=(\psi_1,\ldots,\psi_{N_{h,\Gamma_D}})^T$ will now also include basis functions associated to nodes on Γ_N and Γ_R.

System (33) is solved in a different way if the Dirichlet b.c. are homogeneous ($h_D=0$) or not ($h_D\neq 0$).

If the Dirichlet b.c. are homogeneous or there are no Dirichlet b.c. (i.e., Γ_D is empty), the discretization of the variational formulation associated to (33) is

$$\begin{cases} a(\hat{f}_h,\psi_h)-\int_\Omega \hat{g}_h\psi_h=\int_\Omega u\psi_h+\int_{\Gamma_N}h_N\psi_h+\int_{\Gamma_R}h_R\psi_h \\ \lambda a(\varphi_h,\hat{g}_h)+\sum_{i=1}^n\hat{f}_h(\mathbf{p}_i)\varphi_h(\mathbf{p}_i)=\sum_{i=1}^n z_i\varphi_h(\mathbf{p}_i) \end{cases} \tag{35}$$

for all $\psi_h,\varphi_h\in V^r_{h,\Gamma_D}$, where the bilinear form $a(\cdot,\cdot)$ is now defined as

$$a(\hat{f},\psi)=\int_\Omega\left(\mathbf{K}\nabla\hat{f}\cdot\nabla\psi+\mathbf{b}\cdot\nabla\hat{f}\psi+c\hat{f}\psi\right)+\int_{\Gamma_R}\gamma\hat{f}\psi. \tag{36}$$

Consider the matrices (13)-(15) and the matrix $\boldsymbol{\Psi}$ of basis evaluations (17), with now $\boldsymbol{\psi}=(\psi_1,\ldots,\psi_{N_{h,\Gamma_D}})^T$, and define the matrix

$$\mathbf{B}_R(\gamma)=\int_{\Gamma_R}\gamma\boldsymbol{\psi}\boldsymbol{\psi}^T \tag{37}$$

that represents the Robin b.c.. Using this notation, the Finite Element matrix associated to the bilinear form $a(\cdot,\cdot)$ in (36) is given by

$$\mathbf{A}(\mathbf{K},\mathbf{b},c)=\mathbf{R}_{xx}(\mathbf{K})+\mathbf{R}_{xy}(\mathbf{K})+\mathbf{R}_{yy}(\mathbf{K})+\mathbf{R}_x(\mathbf{b})+\mathbf{R}_y(\mathbf{b})+\mathbf{R}(c)+\mathbf{B}_R(\gamma). \tag{38}$$

We moreover define the vectors

$$(\mathbf{h}_N)_j=\int_{\Gamma_N}h_N\psi_j,\quad (\mathbf{h}_R)_j=\int_{\Gamma_R}h_R\psi_j. \tag{39}$$

Proposition 5. *The Finite Element solution $\hat{f}_h$, when the Dirichlet b.c. are homogeneous or when there are no Dirichlet b.c., exists, is unique and is given by $\hat{f}_h=\boldsymbol{\psi}^T\hat{\mathbf{f}}$ where $\hat{\mathbf{f}}$ is the solution of the linear system*

$$\begin{bmatrix}\boldsymbol{\Psi}^T\boldsymbol{\Psi} & \lambda\mathbf{A}^T \\ \mathbf{A} & -\mathbf{R}\end{bmatrix}\begin{bmatrix}\hat{\mathbf{f}} \\ \hat{\mathbf{g}}\end{bmatrix}=\begin{bmatrix}\boldsymbol{\Psi}^T\mathbf{z} \\ \mathbf{u}+\mathbf{h}_N+\mathbf{h}_R\end{bmatrix} \tag{40}$$

The proof of well-posedness of the discrete problem is given in [1].

It should be noticed that the bilinear form in (36) differs from the one in (11) only for the term corresponding to Robin b.c., and the same of course holds for the Finite Element matrix (38) vs (16). All the non-homogeneous b.c. are instead included in the forcing term of the linear system. For this reason the smoothing matrix $\mathbf{S}$ depends only on the Robin b.c., while the vector $\mathbf{r}$ depends on all the non-homogeneous conditions.

If there are instead non-homogeneous Dirichlet conditions (Γ_D is non-empty and $h_D \neq 0$) we need to define a so-called lifting of the boundary conditions. We consider in this case the space V_h^r in (10) and denote by $\psi_1^D, \ldots, \psi_{N_h^D}^D$ the basis functions associated to the nodes $\boldsymbol{\xi}_1^D, \ldots \boldsymbol{\xi}_{N_h^D}^D$ on Γ_D. The problem is treated in this case by splitting the discrete surface estimator in two parts $f_{D,h}$ and $\hat{s}_h$, with $\hat{f}_h = f_{D,h} + \hat{s}_h$. The first part $f_{D,h} \in \mathrm{span}\{\psi_1^D, \ldots, \psi_{N_h^D}^D\}$ satisfies the non-homogeneous Dirichlet conditions on Γ_D, i.e., $f_{D,h}(\boldsymbol{\xi}_i^D) = h_D(\boldsymbol{\xi}_i^D)$ for $i = 1, \ldots, N_h^D$. The second part $\hat{s}_h \in V_{h,\Gamma_D}^r$ is instead the solution, with homogeneous Dirichlet b.c., of the variational problem:

$$\begin{cases} a(\hat{s}_h, \psi_h) - \int_\Omega \hat{r}_h \psi_h = \int_\Omega u \psi_h + \int_{\Gamma_N} h_N \psi_h + \int_{\Gamma_R} h_R \psi_h - a(f_{D,h}, \psi_h) \\ \lambda a(\varphi_h, \hat{r}_h) + \sum_{i=1}^n \hat{s}_h(\mathbf{p}_i) \varphi_h(\mathbf{p}_i) = \sum_{i=1}^n (z_i - f_{D,h}(\mathbf{p}_i)) \varphi_h(\mathbf{p}_i) \end{cases}$$

for all $\psi_h, \varphi_h \in V_{h,\Gamma_D}^r$, where $\hat{r}_h$ is the adjoint variable associated to $\hat{s}_h$. This system has an extra forcing term, with respect to system (35), that implicitly involves the Dirichlet b.c. h_D through the quantity $f_{D,h}$.

Finally, we can analogously proceed in the areal data framework. See [1] for details.

References

[1] L. Azzimonti, F. Nobile, L.M. Sangalli, and P. Secchi. Mixed finite elements for spatial regression with PDE penalization. Technical Report 20/2013, MOX - Dipartimento di Matematica, Politecnico di Milano, 2013. available at http://mox.polimi.it/it/progetti/pubblicazioni/.

[2] A. Buja, T.J. Hastie, and R.J. Tibshirani. Linear smoothers and additive models. *Annals of Statististics*, (17):153–555, 1989.

[3] B. Ettinger, S. Perotto, and L.M. Sangalli. Spatial regression models over two-dimensional manifolds. Technical Report 54/2012, MOX - Dipartimento di Matematica, Politecnico di Milano, 2012. available at http://mox.polimi.it/it/progetti/pubblicazioni/.

[4] L.C. Evans. *Partial Differential Equations*. American Mathematical Society, 1998.

[5] L. Formaggia, A. Quarteroni, and A. Veneziani. *Cardiovascular mathematics: modeling and simulation of the circulatory system*. Springer, 2009.

[6] S. Guillas and M.J. Lai. Bivariate splines for spatial functional regression models. *J. Nonparametric Stat*, (22):477–497, 2010.

[7] T.J. Hastie and R.J. Tibshirani. *Generalized additive models*. London: Chapman and Hall Ltd., 1990.

[8] F. Lindgren, H. Rue, , and J. Lindström. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, (73):423–498, 2011.

[9] J.L. Lions. *Optimal control of systems governed by partial differential equations*. Springer, 1971.

[10] K.R. Moyle, L. Antiga, and D.A. Steinman. Inlet conditions for image-based CFD models of the carotid bifurcation: is it reasonable to assume fully developed flow? *J Biomech Eng.*, (128):371–379, 2006.

[11] O. Pironneau, F. Hecht, A. Le Hyaric, and J. Morice. *FreeFem++ Software version 3.16*, 2011. Available at www.freefem.org.

[12] A. Quarteroni, R. Sacco, and F. Saleri. *Numerical mathematics*. Springer, 2007.

[13] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, 2012. Available at www.R-project.org.

[14] J.O. Ramsay, G. Hooker, D. Campbell, and J. Cao. Parameter estimation for differential equations: a generalized smoothing approach. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, (69):741–796, 2007.

[15] T. Ramsay. Spline smoothing over difficult regions. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, (64):307–319, 2002.

[16] L.M. Sangalli, J.O. Ramsay, and T. Ramsay. Spatial spline regression models. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, (75,4):1–23, 2013.

[17] A. Veneziani and C. Vergara. Flow rate defective boundary conditions in haemodynamics simulations. *Int. Journ. Num. Meth. Fluids*, (47):803–816, 2005.

[18] S.N. Wood. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, (73):3–36, 2011.

[19] S.N. Wood, M.V. Bravington, and S.L. Hedley. Soap film smoothing. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, (70):931–955, 2008.

MOX Technical Reports, last issues

Dipartimento di Matematica "F. Brioschi",
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

19/2013 Azzimonti, L.; Sangalli, L.M.; Secchi, P.; Domanin, M.; Nobile, F.
Blood flow velocity field estimation via spatial regression with PDE penalization

18/2013 Discacciati, M.; Gervasio, P.; Quarteroni, A.
Interface Control Domain Decomposition (ICDD) Methods for Coupled Diffusion and Advection-Diffusion Problems

17/2013 Chen, P.; Quarteroni, A.
Accurate and efficient evaluation of failure probability for partial different equations with random input data

16/2013 Faggiano, E. ; Lorenzi, T. ; Quarteroni, A.
Metal Artifact Reduction in Computed Tomography Images by Variational Inpainting Methods

15/2013 Antonietti, P.F.; Giani, S.; Houston, P.
Domain Decomposition Preconditioners for Discontinuous Galerkin Methods for Elliptic Problems on Complicated Domains

14/2013 Gianni Arioli, Filippo Gazzola
A new mathematical explanation of the Tacoma Narrows Bridge collapse

13/2013 Pini, A.; Vantini, S.
The Interval Testing Procedure: Inference for Functional Data Controlling the Family Wise Error Rate on Intervals.

12/2013 Antonietti, P.F.; Beirao da Veiga, L.; Bigoni, N.; Verani, M.
Mimetic finite differences for nonlinear and control problems

11/2013 Discacciati, M.; Gervasio, P.; Quarteroni, A.
The Interface Control Domain Decomposition (ICDD) Method for Elliptic Problems

10/2013 Antonietti, P.F.; Beirao da Veiga, L.; Mora, D.; Verani, M.

A stream virtual element formulation of the Stokes problem on polygonal meshes