POLITECNICO DI MILANO
DEPARTMENT OF MATHEMATICS
DOCTORAL PROGRAMME IN MATHEMATICAL MODELS AND METHODS IN ENGENEERING

BLOOD FLOW VELOCITY FIELD ESTIMATION VIA SPATIAL REGRESSION WITH PDE PENALIZATION

Doctoral Dissertation of:
Laura Azzimonti

Supervisor:
Prof. Piercesare Secchi

Co-supervisor:
Dr. Laura M. Sangalli

Tutor:
Prof. Piercesare Secchi

The Chair of the Doctoral Program:
Prof. Roberto Lucchetti

2013 – XXV

Acknowledgments

THIS thesis has been developed within the project MACAREN@MOX, *MAthematichs for CARotid ENdarterectomy @ MOX*, and has been partly supported by MIUR *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/ sangalli/firb.html). Part of the thesis has been developed during a visiting period at MATHICSE-CSQI, École Polytechnique Fédérale de Lausanne, under the supervision of Prof. Fabio Nobile.

I would like to thank my supervisor Piercesare Secchi for his open-minded vision of statistics and for giving me the opportunity to pursue this new branch of mathematics at the interface between statistics and numerical analysis.

I thank my co-supervisor Laura Sangalli for helping me taking the first steps into research and introducing me to spatial functional data analysis. I really appreciate the support she gave me during my studies.

I am very grateful to Fabio Nobile for his time and dedication to my work, helping me making the mathematical details rigorous. The time with him at EPFL was really productive, stimulating and critical to the development of this work.

I would also like to thank Maurizio Domanin for proposing this interesting problem, dedicating his time to this project and providing quality data for us to analyze.

In addition I 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 eco-doppler data.

I thank all the MOXstat group and my PhD colleagues for the interesting discussion and the time spent together. A special thank goes to Stefano, Davide, Alessandro, Francesca, Simone and Bree for the mutual support in the last three years.

Finally, I would like to thank my family who has always encouraged me.

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 method is also extended to model dynamic surfaces evolving in time. 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.

Contents

Introduction

In this thesis we propose a novel non-parametric regression technique for surfaces and spatial fields estimation, able to include a prior knowledge on the shape of the surface or field and to comply with complex conditions at the boundary of the problem domain. The proposed technique is particularly well suited for applications in physics, engineering, biomedicine, etc., where a prior knowledge on the bidimensional or three-dimensional field might be available from physical principles and should be taken into account in the field estimation or smoothing process. We consider in particular phenomena where the field can be described by a partial differential equation (PDE) and has to satisfy some known boundary conditions. Specifically we focus on phenomena that are well described by linear second order elliptic PDEs, typically transport-reaction-diffusion problems.

Spatial Regression with PDE penalization (SR-PDE) uses a functional data analysis approach (see, e.g., [30] and [13]) and generalizes classical spatial smoothing techniques, such as thin-plate splines. We propose, in fact, to estimate the surface or the field minimizing a penalized least square functional, with the roughness penalty involving the PDE governing the phenomenon. The proposed methodology allows for important modeling flexibility, accounting for space anisotropy and non-stationarity in a straightforward way, as well as unidirectional smoothing effects. SR-PDE has very broad applicability since PDEs are commonly used to describe phenomena behavior in many fields of physics, mechanics, biology and engineering. In particular 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* (*MACAREN@MOX*), which 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.

Many methods for surface estimation define the estimate as the minimizer of a penalized sum-of-square-error functional, with the penalty term involving a simple partial differential operator. Thin-plate spline smoothing, for example, penalizes an energy functional in $\mathbb{R}^2$ that involves second order derivatives, i.e.,

$$\int_{\mathbb{R}^2} \left(\frac{\partial^2 f}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 f}{\partial y^2}\right)^2 . \tag{1}$$

The minimizer of this functional belongs to the linear space generated by the Green's functions associated to the bilaplacian (see [36] for details). Thin-plate spline smoothing has been first extended to the case of bounded domains in [33], where the thin-plate energy (1) is computed only over the bounded domain of interest. Since the minimizer cannot be directly characterized, it is approximated by a surface in the space of tensor product B-splines.
Recently, more complex smoothing methods have been developed, dealing with general bounded domains in $\mathbb{R}^2$ and general boundary conditions. Some examples are the spatial functional regression models introduced in [17], Soap-Film Smoothing (SOAP), described in [38] and extended to the time dimension in [1, 24], and the Spatial Spline Regression models (SSR) described in [32], which generalize the Finite Element L-splines introduced in [31]. All these methods estimate bidimensional surfaces on complex bounded domains penalizing a simple differential operator. The spatial models proposed in [17], in particular, consider a regularizing term that involves a linear combination of all the partial derivatives up to a chosen order r. These models employ bivariate splines over triangulations (see, e.g., [21]), which provide a basis for piecewise-polynomial surfaces, to solve the estimation problem.
On the other hand both soap-film smoothing and spatial spline regression models estimate bidimensional surfaces penalizing, on the domain of interest Ω, the Laplace operator of the surface as a measure of the local curvature, i.e.,

$$\int_{\Omega} \left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}\right)^2 .$$

Soap-film smoothing approximates the minimizer of the penalized least square functional with a linear combination of Green's functions of the bilaplacian on the domain of interest, centered on the vertices of a lattice. Finite Element L-splines and spatial spline regression models solve instead directly the PDE associated to the penalized least square functional by means of the Finite Element method, which provides a basis for piecewise polynomial surfaces.
Finally, SR-PDE have also strong connections with the method proposed in [22] where Gaussian fields and Gaussian Markov random fields are linked via a stochastic partial differential equation that induces a Matérn covariance. This method resort to the Finite Element method to solve partial differential equations over irregular grids of points.

Following the approach presented in [31] and [32], we propose to estimate the field minimizing a least square functional penalized with a roughness term that involves, instead of the simple Laplacian, a more general PDE modeling the phenomenon $Lf = u$, i.e.,

$$\int_{\Omega} (Lf - u)^2 .$$

This approach is similar to the one used in control theory when a distributed control is considered; see for example [23].

Likewise in [31] and [32], SR-PDE exploits advanced numerical analysis techniques and, specifically, it makes use of the Finite Element method. In particular for the discretization of the estimation problem we resort to a mixed equal order Finite Element method, similar to classical mixed methods for the discretization of fourth order problems (see, e.g., [6]). The proposed mixed method provides a good approximation of the field but also of its first and second order derivatives that can be useful in order to compute physical quantities of interest. The resulting estimators have a typical penalized regression form, they are linear in the observed data values and classical inferential tools can be derived. SR-PDE is currently implemented in R [28] and in FreeFem++ [26].

The smoothing technique is also extended to the case of areal data, i.e., data that represent quantities computed on some subdomains, particularly interesting in many applications and specifically in the velocity field estimation problem from echo-doppler acquisitions. The data represent, in fact, the mean velocity of blood computed on some subdomains located on the considered artery section.

Finally the time dependence could also be introduced in SR-PDE models, in order to deal with data depending both on space and time. This extension allows to estimate time evolving surfaces, particularly interesting in the velocity field application in order to study how the velocity field evolves during a heartbeat. To estimate dynamic surfaces we propose to minimize a least square functional penalized with the misfit of a time dependent PDE that models the phenomenon of interest.

In Chapter 1 the motivating problem leading the study is detailed and SR-PDE is described. The model for pointwise and areal data is derived and the surface estimator is obtained in both the frameworks. The discretization of the surface estimator by means of the mixed Finite Element method is introduced and, thanks to its linearity, the distributional properties of the estimator are derived. SR-PDE is then compared to standard SSR and SOAP in different simulation settings with data distributed uniformly on the domain or only on some subregions. The simulation studies show that the inclusion of the prior knowledge on the shape of the surface improves significantly its estimate. Finally the results obtained in the velocity field estimation using SR-PDE smoothing are shown.

In Chapter 2 the technical assumptions made in SR-PDE models and the proofs of the existence and uniqueness of the surface estimator are detailed. The mixed Finite Element method for the discretization of the estimation problem is derived and its well posedness is proved. The mean of the Finite Element estimator is proved to be convergent, with an optimal convergence rate, to the true underlying surface when the characteristic size associated to the Finite Element basis goes to zero. The theoretical results are confirmed by numerical experiments. The properties of the estimator in the areal setting are obtained along the same line followed in the pointwise framework.

In Chapter 3 SR-PDE models are extended in order to deal with data dependent on space and time. The results of existence and uniqueness of the surface estimator, presented in Chapters 1 and 2, are extended to the case of time dependent PDEs. The discretization of the dynamic surface estimator by means of the Finite Element method and the Finite Difference method is derived. Finally the results obtained in a toy example and in the velocity field estimation using time-dependent SR-PDE are shown.

In Chapter 3.7 future research directions are presented.

CHAPTER 1

Spatial regression with PDE penalization

1.1 Introduction

In this chapter we describe Spatial Regression with PDE penalization (SR-PDE) that is a non-parametric regression technique for surface estimation on bounded and complex domains. This method includes a prior knowledge on the shape of the surface described in terms of a PDE governing the phenomenon under study.

This technique has been developed within the project *MACAREN@MOX* in order to estimate, from echo-color doppler acquisitions, the velocity field of the blood on a cross-section of the common carotid artery. The anatomy of the carotid artery is shown in Figure 1.1. The *MACAREN@MOX* project gathers researchers from different mathematical fields and medical doctors in cardiac surgery with the aim of investigating the pathogenesis of atherosclerosis in human carotids. The project intends specifically to

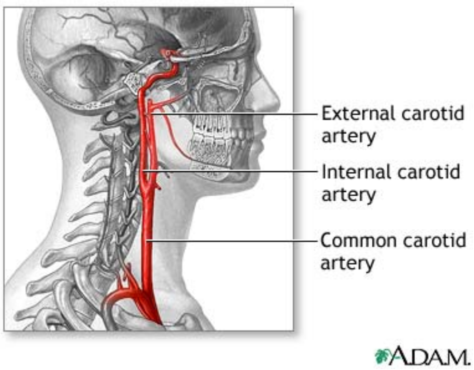

Figure 1.1: *Anatomy of the carotid artery.*

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 [25] via numerical simulations of the blood flow on real patient-specific vessel morphologies.

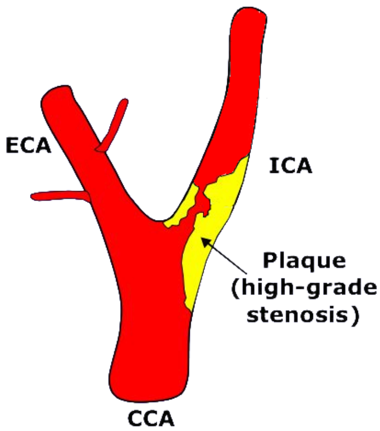

Figure 1.2: *High grade stenosis (>70%) in the internal carotid artery.*

The data collected within the project include: Echo-Color Doppler (ECD) measurements of blood flow at a cross-section of the common carotid artery (CCA), 2 cm before the carotid bifurcation, for patients affected by high grade stenosis (>70%) in the internal carotid artery (ICA), as shown in Figure 1.2; 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.

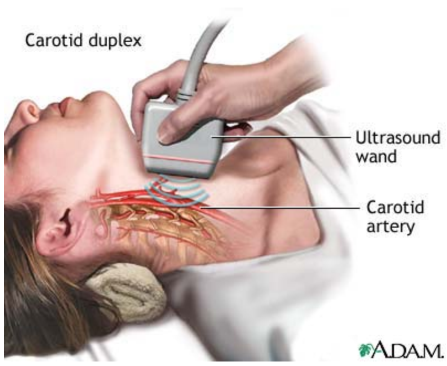

Figure 1.3: *Carotid echo-color doppler scan.*

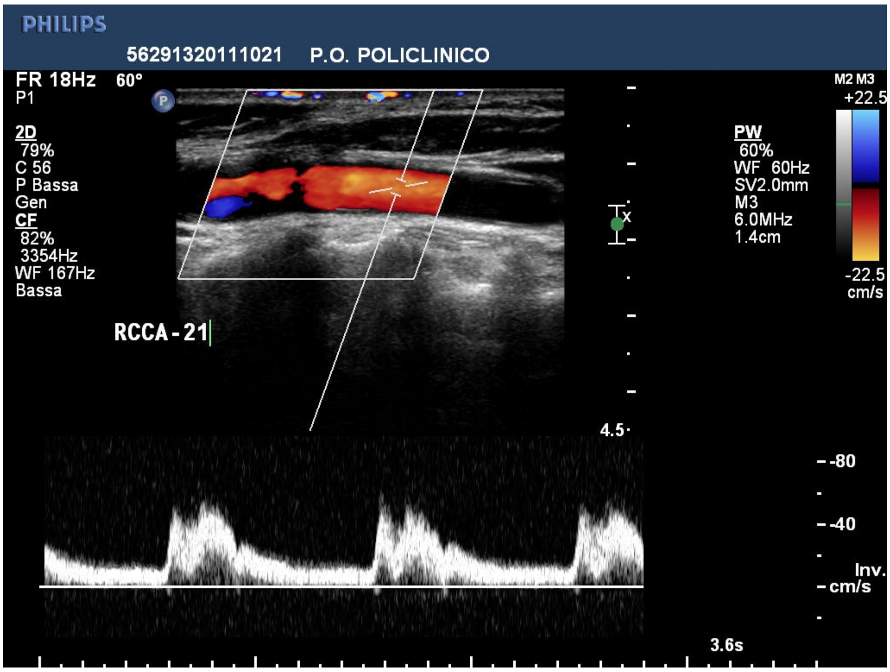

Figure 1.4: *Echo-color doppler image corresponding to the central point of the carotid section located 2 cm before the carotid bifurcation.*

Carotid Echo-Color Doppler is a medical imaging procedure, shown in Figure 1.3, 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. Figure 1.4 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

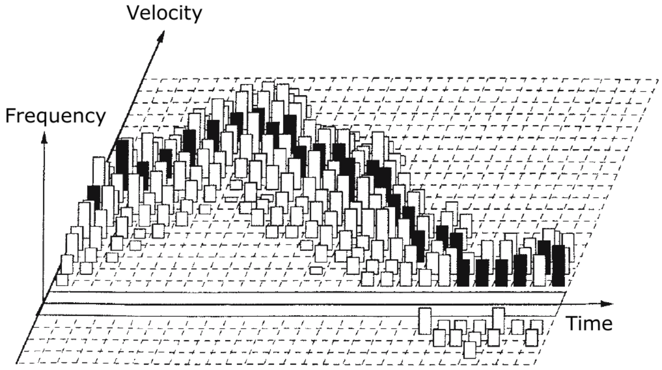

Figure 1.5: *Graphical display of the echo doppler signal.*

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, as shown in Figure 1.5, 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 thesis, we shall consider a fixed time instant corresponding to the systolic peak, which is of crucial clinical interest.

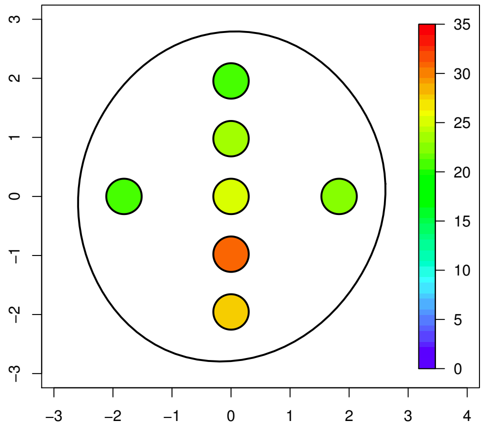

Figure 1.6: *MRI reconstruction of the cross-section of the carotid artery located 2 cm before the bifurcation; cross-shaped pattern of observations with each beam colored according to the mean blood-velocity measured on the beam at systolic peak time.*

Figure 1.6 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 we have a 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 [14] and references therein. Firstly, the physics of the problem implies that the blood-flow velocity is zero at the arterial wall, due to the friction between blood cells and the arterial wall; these conditions, called no-slip boundary conditions, have to be imposed in the surface estimation in order to obtain physiological results. For what concerns the shape of the surface, 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.

Classical non-parametric 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 [31] account explicitly for the shape of the domain, efficiently dealing with irregular shaped domains; soap-film smoothing (SOAP), described in [38], considers both the shape of the domain and some common types of boundary conditions; Spatial Spline Regression (SSR), presented in [32], extends [31] and includes general boundary conditions. The methods in [31], [38] and [32] 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. For this reason, extending [31] and [32], we develop a non-parametric method that includes in the model the prior information on the phenomenon under study, formalized in terms of a governing PDE. Specifically, SR-PDE features a roughness term that involves, instead of the simple Laplacian, a more general PDE modeling the phenomenon.

The chapter is organized as follows. Section 1.2 introduces SR-PDE for pointwise observations. Section 1.3 extends the models to the case of areal data, which is of interest in the analysis of ECD measurements here considered. Section 1.4 describes the Finite Element solution to the estimation problem and derives the inferential properties of the estimators. Section 1.5 deals with general boundary conditions. Section 1.6 compares SR-PDE 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. Section 1.7 presents the application within the MACAREN@MOX project: details on the ECD acquisitions are given and the results obtained with SR-PDE are shown.

1.2 Spatial Regression with PDE penalization 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

$$z_i = f_0(\mathbf{p}_i) + \epsilon_i \tag{1.1}$$

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 a 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, the 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 [31] and [32], we propose to estimate f_0 by minimizing the penalized sum-of-square-error functional

$$J(f) = \sum_{i=1}^{n} \left(f(\mathbf{p}_i) - z_i\right)^2 + \lambda \int_{\Omega} \left(Lf(\mathbf{p}) - u(\mathbf{x})\right)^2 d\mathbf{x} \tag{1.2}$$

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 (div∇f), first order differential operators as the gradient (∇f) and also the identity (f); the general form that we consider is

$$Lf = -\text{div}(\mathbf{K}\nabla f) + \mathbf{b} \cdot \nabla f + cf \tag{1.3}$$

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 [31] and [32], where the Laplacian Δf is penalized, thus controlling the local curvature of f.

The three terms that compose the general second order operator (1.3) provide different smoothing effects. The diffusion term $-\text{div}(\mathbf{K}\nabla f)$ 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

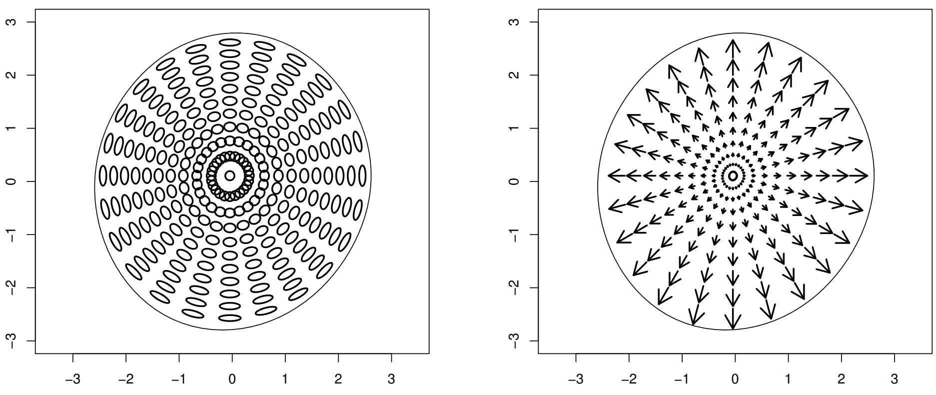

Figure 1.7: *Left: diffusion tensor field* $\mathbf{K}$*, used in the velocity field application, that smooths the observations in the tangential direction of concentric circles. Right: transport field* $\mathbf{b}$*, used in the velocity field application, that smooths the observations in the radial direction, from the center of the section to the boundary.*

space varying anisotropic diffusion tensor that smooths the observations in the tangential direction of concentric circles (see Figure 1.7 Left); a transport field that smooths the observations in the radial direction, from the center of the section to the boundary (see Figure 1.7 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 [12] 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 $\partial\Omega$ 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 1.5 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 thesis 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 $\mathbf{x}$; unless differently specified, the integrals are computed with respect of the Lebesgue measure, i.e., $\int_D q = \int_D q(\mathbf{x})d\mathbf{x}$, 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 [32].

1.2.1 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:*

$$\begin{cases} L\hat{f} = u + \hat{g} & in\ \Omega \\ \hat{f} = 0 & on\ \partial\Omega \end{cases} \qquad \begin{cases} L^*\hat{g} = -\frac{1}{\lambda}\sum_{i=1}^{n}(\hat{f} - z_i)\delta_{\mathbf{p}_i} & in\ \Omega \\ \hat{g} = 0 & on\ \partial\Omega \end{cases} \tag{1.4}$$

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*

$$L^*\hat{g} = -div(\mathbf{K}\nabla\hat{g}) - \mathbf{b}\cdot\nabla\hat{g} + (c - div(\mathbf{b}))\hat{g}. \tag{1.5}$$

The proof of Proposition 1 is based on PDE optimal control theory (see, e.g., [23]) and is detailed in Chapter 2, 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 1.5.

1.3 Spatial Regression with PDE penalization 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 $i = 1, \ldots, N$, be some subdomains where we have observations and z_{ij}, for $j = 1, \ldots, n_i$, be the observations located at point $\mathbf{p}_{ij} \in D_i$. For the observations z_{ij}, we consider the pointwise model (1.1), i.e.,

$$z_{ij} = f_0(\mathbf{p}_{ij}) + \epsilon_{ij} \tag{1.6}$$

where ϵ_{ij}, for $i = 1, \ldots, N$ and $j = 1, \ldots, n_i$, 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 1.4) 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 $i = 1, \ldots, N$. From (1.6), 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 $\zeta_i = \sum_{j=1}^{n_i} \epsilon_{ij}/n_i$, $i = 1, \ldots, N$, 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 $\mathbb{E}\left[f_0(P) | P \in D_i\right]$ 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.,

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

We may thus consider the following model:

$$\bar{z}_i = \frac{1}{|D_i|} \int_{D_i} f_0 + \eta_i \tag{1.7}$$

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

$$\bar{J}(f) = \sum_{i=1}^{N} \frac{1}{|D_i|} \left(\int_{D_i} (f - \bar{z}_i) \right)^2 + \lambda \int_{\Omega} (Lf - u)^2 \tag{1.8}$$

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 (1.8) 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.

1.3.1 Solution to the estimation problem

The estimation problem can be formulated as follows.

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

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

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:*

$$\begin{cases} L\hat{f} = u + \hat{g} & \text{in } \Omega \\ \hat{f} = 0 & \text{on } \partial\Omega \end{cases} \quad \begin{cases} L^*\hat{g} = -\frac{1}{\lambda}\sum_{i=1}^{N} \frac{1}{|D_i|}\mathbb{I}_{D_i}\int_{D_i}(\hat{f} - \bar{z}_i) & \text{in } \Omega \\ \hat{g} = 0 & \text{on } \partial\Omega \end{cases} \tag{1.9}$$

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 Chapter 2.

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 $\mathbb{E}\left[f_0(P)|\, P \in D_i\right]$:*

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

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

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

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

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

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 $\mathbb{V}ar(\bar{z}_i) \propto 1/\mu(D_i)$.

1.4 Finite Element estimator and its distributional properties

The surface estimation problems in the pointwise and areal data frameworks presented respectively in Sections 1.2 and 1.3 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., [27]), 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 1.8 Left shows the triangulation considered in the velocity field application. We

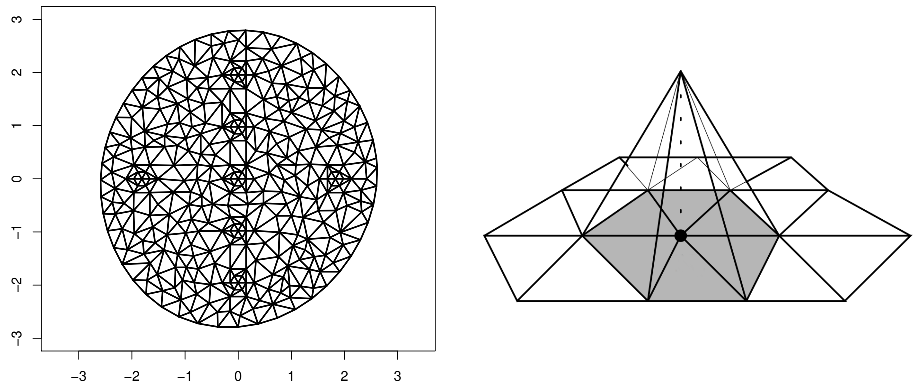

Figure 1.8: *Left: triangulation of the carotid cross-section of interest in the velocity field application. Right: a linear Finite Element basis function on a triangulation.*

consider the space V_h^r of piecewise continuous polynomial functions of order $r \geq 1$ over the triangulation:

$$V_h^r = \left\{ v \in C^0(\bar{\Omega}) : v|_K \in \mathbb{P}^r(K) \, \forall K \in \mathcal{T}_h \right\}. \tag{1.10}$$

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 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 1.8 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

$$V_{h,0}^r = \{v \in C^0(\bar{\Omega}) : v|_{\partial\Omega} = 0 \text{ and } v|_K \in \mathbb{P}^r(K) \forall K \in \mathcal{T}_h\} \tag{1.11}$$

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 1.5 we extend the results presented in this section to the case of more general boundary conditions.

1.4.1 Pointwise estimator

In order to define the weak problem associated to the system of PDEs (1.4) we introduce the bilinear form $a(\cdot,\cdot)$ associated to the operator L, defined as

$$a(\hat{f},\psi) = \int_\Omega \left(\mathbf{K}\nabla\hat{f}\cdot\nabla\psi + \mathbf{b}\cdot\nabla\hat{f}\psi + c\hat{f}\psi\right). \tag{1.12}$$

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

$$\begin{cases} a(\hat{f}_h,\psi_h) - \int_\Omega \hat{g}_h\psi_h = \int_\Omega u\psi_h \\ \lambda a(\varphi_h,\hat{g}_h) + \sum_{i=1}^n \hat{f}_h(\mathbf{p}_i)\varphi_h(\mathbf{p}_i) = \sum_{i=1}^n z_i\varphi_h(\mathbf{p}_i) \end{cases} \tag{1.13}$$

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

$$\mathbf{R}(c) = \int_\Omega c\boldsymbol{\psi}\boldsymbol{\psi}^T, \quad \mathbf{R}_x(\mathbf{b}) = \int_\Omega \mathbf{b}_1\boldsymbol{\psi}\boldsymbol{\psi}_x^T, \quad \mathbf{R}_y(\mathbf{b}) = \int_\Omega \mathbf{b}_2\boldsymbol{\psi}\boldsymbol{\psi}_y^T, \tag{1.14}$$

$$\mathbf{R}_{xx}(\mathbf{K}) = \int_\Omega \mathbf{K}_{11}\boldsymbol{\psi}_x\boldsymbol{\psi}_x^T, \quad \mathbf{R}_{yy}(\mathbf{K}) = \int_\Omega \mathbf{K}_{22}\boldsymbol{\psi}_y\boldsymbol{\psi}_y^T, \tag{1.15}$$

$$\mathbf{R}_{xy}(\mathbf{K}) = \int_\Omega \mathbf{K}_{12}(\boldsymbol{\psi}_x\boldsymbol{\psi}_y^T + \boldsymbol{\psi}_y\boldsymbol{\psi}_x^T), \tag{1.16}$$

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 (1.12) is given by

$$\mathbf{A}(\mathbf{K},\mathbf{b},c) = \mathbf{R}_{xx}(\mathbf{K}) + \mathbf{R}_{xy}(\mathbf{K}) + \mathbf{R}_{yy}(\mathbf{K}) + \mathbf{R}_x(\mathbf{b}) + \mathbf{R}_y(\mathbf{b}) + \mathbf{R}(c). \tag{1.17}$$

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

$$\mathbf{R} = \mathbf{R}(1) = \int_\Omega \boldsymbol{\psi}\boldsymbol{\psi}^T \tag{1.18}$$

and

$$\boldsymbol{\Psi} = \begin{bmatrix} \boldsymbol{\psi}^T(\mathbf{p}_1) \\ \vdots \\ \boldsymbol{\psi}^T(\mathbf{p}_n) \end{bmatrix} \tag{1.19}$$

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 (1.13) 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*

$$\begin{bmatrix} \boldsymbol{\Psi}^T\boldsymbol{\Psi} & \lambda\mathbf{A}^T \\ \mathbf{A} & -\mathbf{R} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Psi}^T\mathbf{z} \\ \mathbf{u} \end{bmatrix}. \tag{1.20}$$

The proof of well-posedness of the discrete problem is given in Chapter 2.

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

$$\hat{\mathbf{z}} = \mathbf{S}\mathbf{z} + \mathbf{r} \tag{1.21}$$

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

$$\mathbf{S} = \boldsymbol{\Psi}\left(\boldsymbol{\Psi}^T\boldsymbol{\Psi} + \lambda\mathbf{P}\right)^{-1}\boldsymbol{\Psi}^T, \tag{1.22}$$

$$\mathbf{r} = \boldsymbol{\Psi}\left(\boldsymbol{\Psi}^T\boldsymbol{\Psi} + \lambda\mathbf{P}\right)^{-1}\lambda\mathbf{P}\mathbf{A}^{-1}\mathbf{u}. \tag{1.23}$$

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

$$\mathbf{P} = \mathbf{P}(\mathbf{K}, \mathbf{b}, c) = \mathbf{A}^T(\mathbf{R})^{-1}\mathbf{A}. \tag{1.24}$$

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 (1.2). Notice that, thanks to the weak formulation of the estimation problem, this penalty matrix does not involve the computation of second order derivatives. We can show that, on the Finite Element space used to discretize the problem, $\mathbf{P}$ is analogue to the penalty matrix $\tilde{\mathbf{P}}$ that would be obtained as direct discretization of the penalty term in (1.2), 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.

Remark 2. *Assume for the moment that ψ_j are smooth functions and neglect the forcing term u. The penalty matrix $\mathbf{P}$ in (1.24) 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 (1.2)*, 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

$$\int_\Omega \delta(\mathbf{s},\mathbf{t})q(\mathbf{t})d\mathbf{t} = q(\mathbf{s}) \;\; \forall q \in L^2(\Omega)\cap C(\Omega). \tag{1.25}$$

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* (1.25) *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*$\{\psi_1,\ldots,\psi_{N_{h,0}}\}$.

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 [32]). 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 $\mathbb{C}\text{ov}(\mathbf{z})=\sigma^2\mathbf{I}$ in our model definition, we can compute the expected value and the variance of the estimator $\hat{\mathbf{z}}$:

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

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., [5, 19]). Hence we can estimate σ^2 as

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

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

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

1.4.2 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 the system of PDEs (1.9) can be discretized as

$$\begin{cases} a(\hat{f}_h,\psi_h)-\int_\Omega \hat{g}_h\psi_h=\int_\Omega u\psi_h \\ \lambda a(\varphi_h,\hat{g}_h)+\sum_{i=1}^N\frac{1}{|D_i|}\int_{D_i}\hat{f}_h\int_{D_i}\varphi_h=\sum_{i=1}^N\bar{z}_i\int_{D_i}\varphi_h\end{cases}\tag{1.26}$$

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 (1.12).

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 $\mathbf{W} = \text{diag}(|D_1|, \ldots, |D_N|)$, recalling 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*

$$\begin{bmatrix} \bar{\mathbf{\Psi}}^T \mathbf{W} \bar{\mathbf{\Psi}} & \lambda \mathbf{A}^{\mathbf{T}} \\ \mathbf{A} & -\mathbf{R} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{\Psi}}^T \mathbf{W} \bar{\mathbf{z}} \\ \mathbf{u} \end{bmatrix}. \tag{1.27}$$

The proof of the proposition is detailed in Chapter 2.

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:

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

The Finite Element counterpart of this estimator is defined as

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

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

Properties of the estimator

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

$$\hat{\bar{\mathbf{z}}} = \bar{\mathbf{S}} \bar{\mathbf{z}} + \bar{\mathbf{r}} \tag{1.28}$$

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

$$\bar{\mathbf{S}} = \bar{\mathbf{\Psi}} \left(\bar{\mathbf{\Psi}}^T \mathbf{W} \bar{\mathbf{\Psi}} + \lambda \mathbf{P} \right)^{-1} \bar{\mathbf{\Psi}}^T \mathbf{W}, \tag{1.29}$$

$$\mathbf{b} = \bar{\mathbf{\Psi}} \left(\bar{\mathbf{\Psi}}^T \mathbf{W} \bar{\mathbf{\Psi}} + \lambda \mathbf{P} \right)^{-1} \lambda \mathbf{P} \mathbf{A}^{-1} \mathbf{u}. \tag{1.30}$$

From the definition of model (1.7) and the linearity of the estimator we can derive the mean of the estimator

$$\mathbb{E}\left[\hat{\bar{\mathbf{z}}}\right] = \bar{\mathbf{S}} \bar{\mathbf{z}}_0 + \bar{\mathbf{b}}, \tag{1.31}$$

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

$$\mathbb{C}\mathrm{ov}(\hat{\bar{\mathbf{z}}}) = \bar{\mathbf{S}} \;\mathrm{diag}(\bar{\sigma}_1^2, \ldots, \bar{\sigma}_N^2)\; \bar{\mathbf{S}}^T. \tag{1.32}$$

It should be noticed that in the areal data framework the expected value (1.31) and the variance (1.32) 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 (1.7) provides information only on the areal errors η_i. In particular, in the considered areal framework, the variance

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

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

1.5 General boundary conditions

All the results presented in Sections 1.2, 1.3 and 1.4 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

$$\begin{cases} f = h_D & \text{on } \Gamma_D \\ \mathbf{K}\nabla f \cdot \boldsymbol{\nu} = h_N & \text{on } \Gamma_N \\ \mathbf{K}\nabla f \cdot \boldsymbol{\nu} + \gamma f = h_R & \text{on } \Gamma_R \end{cases} \tag{1.33}$$

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., [12]).
Under (1.33), the solution of the estimation problem and of its discrete counterpart also involve boundary terms. The space V is now the space of functions in $L^2(\Omega)$ with first and second derivatives in $L^2(\Omega)$ that satisfy (1.33).
Starting with the pointwise data framework, the estimation problem with general b.c. (1.33) is analogous to Problem 1. The unique solution of the problem, $\hat{f} \in V$, is obtained by solving

$$\begin{cases} L\hat{f} = u + \hat{g} & \text{in } \Omega \\ + \, b.c. & \text{on } \partial\Omega \end{cases} \qquad \begin{cases} L^*\hat{g} = -\frac{1}{\lambda}\sum_{i=1}^n (\hat{f} - z_i)\delta_{\mathbf{p}_i} & \text{in } \Omega \\ + \, b.c.^* & \text{on } \partial\Omega \end{cases} \tag{1.34}$$

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.,

$$\begin{cases} g = 0 & \text{on } \Gamma_D \\ \mathbf{K}\nabla g \cdot \boldsymbol{\nu} + \mathbf{b} \cdot \boldsymbol{\nu} g = 0 & \text{on } \Gamma_N \\ \mathbf{K}\nabla g \cdot \boldsymbol{\nu} + (\mathbf{b} \cdot \boldsymbol{\nu} + \gamma) g = 0 & \text{on } \Gamma_R. \end{cases} \tag{1.35}$$

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 (1.34) 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 the system of PDEs (1.34) is

$$\begin{cases} a(\hat{f}_h, \psi_h) - \int_\Omega \hat{g}_h \psi_h = \int_\Omega u \psi_h + \int_{\Gamma_N} h_N \psi_h + \int_{\Gamma_R} h_R \psi_h \\ \lambda a(\varphi_h, \hat{g}_h) + \sum_{i=1}^n \hat{f}_h(\mathbf{p}_i) \varphi_h(\mathbf{p}_i) = \sum_{i=1}^n z_i \varphi_h(\mathbf{p}_i) \end{cases} \tag{1.36}$$

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

$$a(\hat{f}, \psi) = \int_\Omega \left(\mathbf{K}\nabla \hat{f} \cdot \nabla \psi + \mathbf{b} \cdot \nabla \hat{f} \psi + c \hat{f} \psi \right) + \int_{\Gamma_R} \gamma \hat{f} \psi. \tag{1.37}$$

Consider the matrices (1.14)-(1.16) and the matrix $\boldsymbol{\Psi}$ of basis evaluations (1.19), with now $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_{N_{h,\Gamma_D}})^T$, and define the matrix

$$\mathbf{B}_R(\gamma) = \int_{\Gamma_R} \gamma \boldsymbol{\psi} \boldsymbol{\psi}^T \tag{1.38}$$

that represents the Robin b.c.. Using this notation, the Finite Element matrix associated to the bilinear form $a(\cdot,\cdot)$ in (1.37) is given by

$$\mathbf{A}(\mathbf{K}, \mathbf{b}, c) = \mathbf{R}_{xx}(\mathbf{K}) + \mathbf{R}_{xy}(\mathbf{K}) + \mathbf{R}_{yy}(\mathbf{K}) + \mathbf{R}_x(\mathbf{b}) + \mathbf{R}_y(\mathbf{b}) + \mathbf{R}(c) + \mathbf{B}_R(\gamma). \tag{1.39}$$

We moreover define the vectors

$$(\mathbf{h}_N)_j = \int_{\Gamma_N} h_N \psi_j, \;\; (\mathbf{h}_R)_j = \int_{\Gamma_R} h_R \psi_j. \tag{1.40}$$

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*

$$\begin{bmatrix} \mathbf{\Psi}^T\mathbf{\Psi} & \lambda \mathbf{A}^T \\ \mathbf{A} & -\mathbf{R} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{bmatrix} = \begin{bmatrix} \mathbf{\Psi}^T \mathbf{z} \\ \mathbf{u} + \mathbf{h}_N + \mathbf{h}_R \end{bmatrix} \tag{1.41}$$

The proof of the proposition is detailed in Chapter 2.

It should be noticed that the bilinear form in (1.37) differs from the one in (1.12) only for the term corresponding to Robin b.c., and the same of course holds for the Finite Element matrix (1.39) vs (1.17). 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 (1.10) 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 $f_{D,h} \in \text{span}\{\psi_1^D, \ldots, \psi_{N_h^D}^D\}$ 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 (1.36), 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 Chapter 2 for details.

1.6 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 1.9, 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 $\mathbf{p}_1, \ldots, \mathbf{p}_n$ uniformly sampled on the entire domain;

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

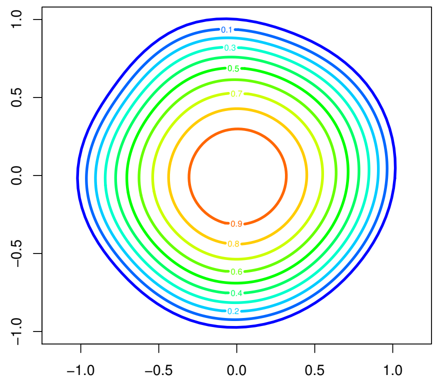

Figure 1.9: *True surface f_0, with almost circular isolines and zero value at the boundary of the domain, used for the simulation studies; the image displays the isolines* $(0, 0.1, \ldots, 0.9, 1)$.

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 (1.1) with the true function f_0 displayed in Figure 1.10.

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 1.10 Left, 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

$$\mathbf{K}(x, y) = \begin{bmatrix} y^2 + \kappa_1 x^2 & (\kappa_1 - 1)xy \\ (\kappa_1 - 1)xy & x^2 + \kappa_1 y^2 \end{bmatrix} + \kappa_2 \left(R^2 - x^2 - y^2\right) \mathbf{I}_2, \tag{1.42}$$

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

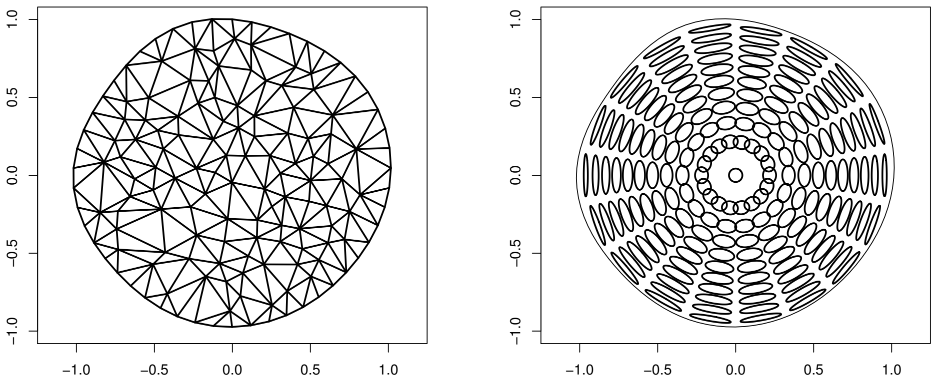

Figure 1.10: *Left: triangulation of the domain Ω used in the simulation studies for SSR and SR-PDE. Right: diffusion tensor field $\mathbf{K}$ used in the simulation studies for SR-PDE.*

of Figure 1.10. 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 (1.42), 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 (1.42), 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 [37], using 49 interior knots on a lattice.

Figures 1.11-1.13 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 1.10.

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

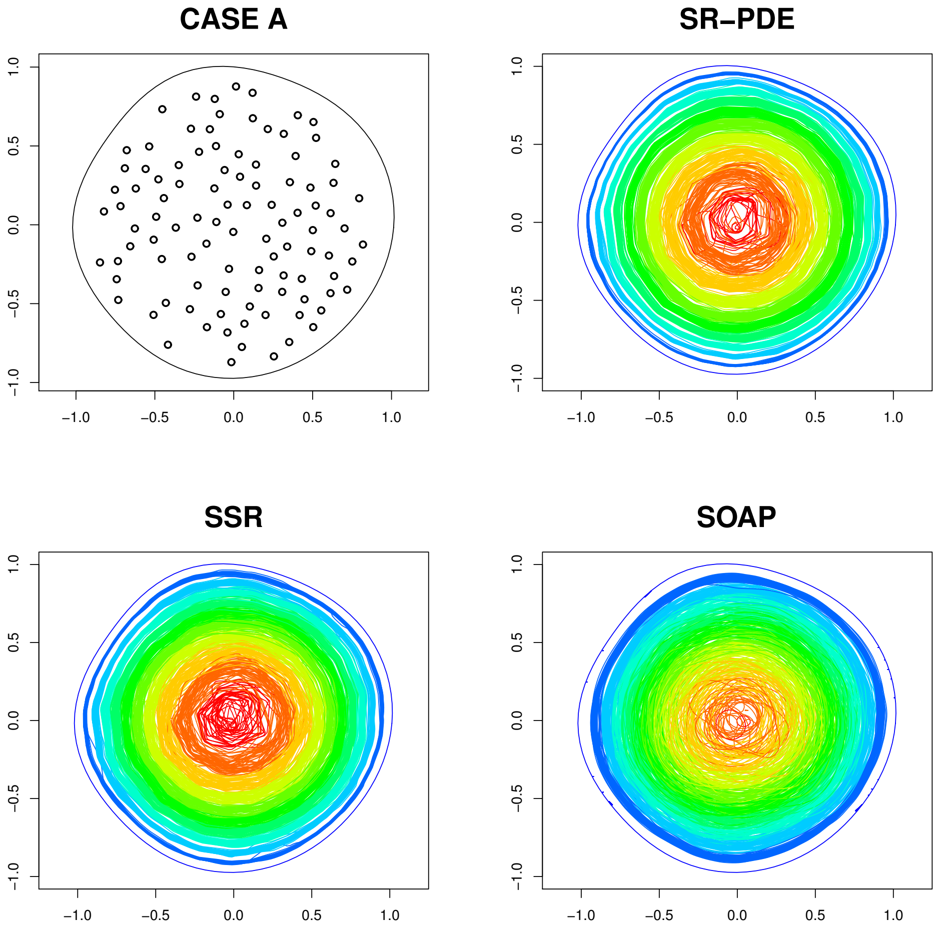

Figure 1.11: *Top left: location points sampled in the first replicate for case A. Top right, bottom left, bottom right: surface estimates obtained using respectively SR-PDE, SSR and SOAP; the images display the isolines* $(0, 0.1, \ldots, 0.9, 1)$ *of the surface estimates obtained in the 50 simulation replicates; the isolines are colored using the same color scale used for the isolines of the true function* f_0 *in Figure 1.10.*

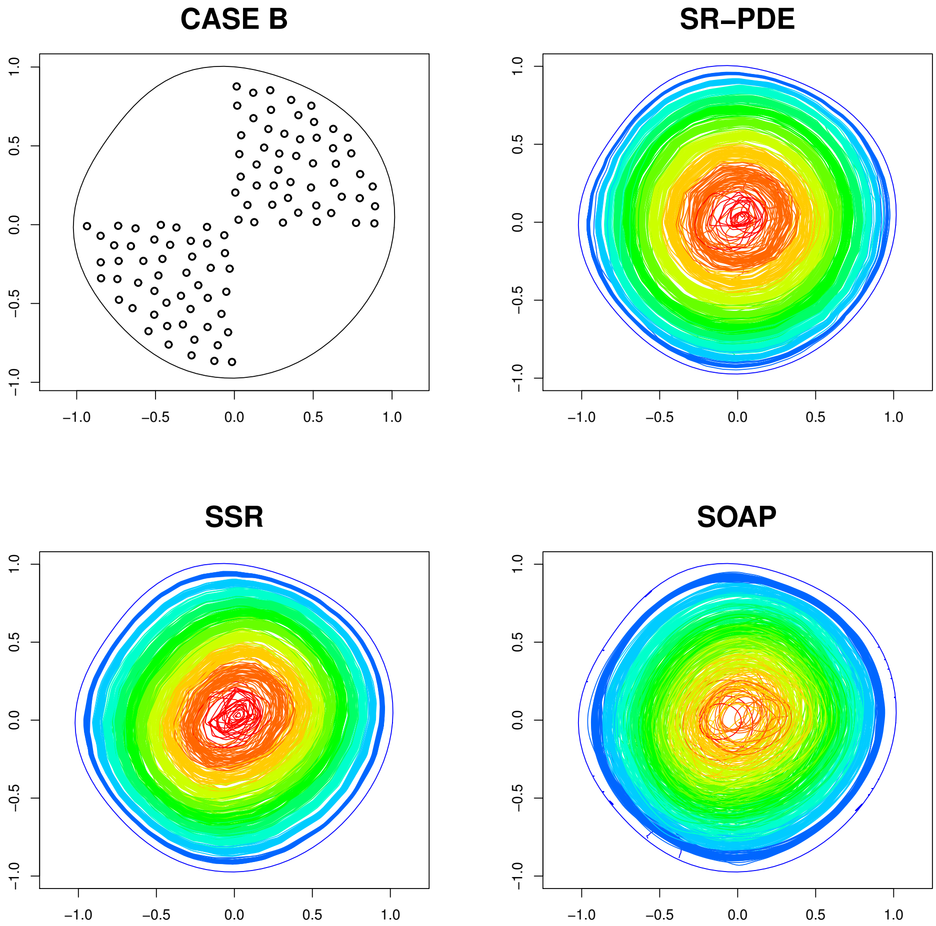

Figure 1.12: *Top left: location points sampled in the first replicate for case B. Top right, bottom left, bottom right: surface estimates obtained using respectively SR-PDE, SSR and SOAP; the images display the isolines* $(0, 0.1, \ldots, 0.9, 1)$ *of the surface estimates obtained in the 50 simulation replicates; the isolines are colored using the same color scale used for the isolines of the true function* f_0 *in Figure 1.10.*

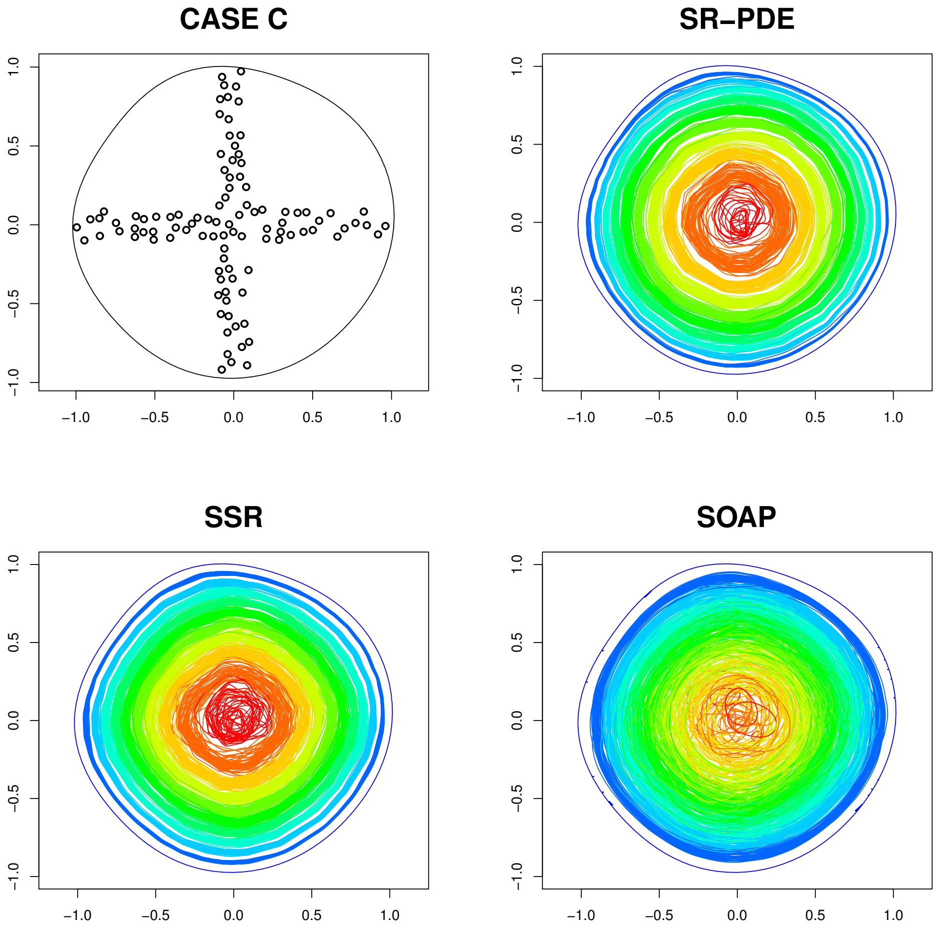

Figure 1.13: *Top left: location points sampled in the first replicate for case C. Top right, bottom left, bottom right: surface estimates obtained using respectively SR-PDE, SSR and SOAP; the images display the isolines* $(0, 0.1, \ldots, 0.9, 1)$ *of the surface estimates obtained in the 50 simulation replicates; the isolines are colored using the same color scale used for the isolines of the true function* f_0 *in Figure 1.10.*

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 1.14 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

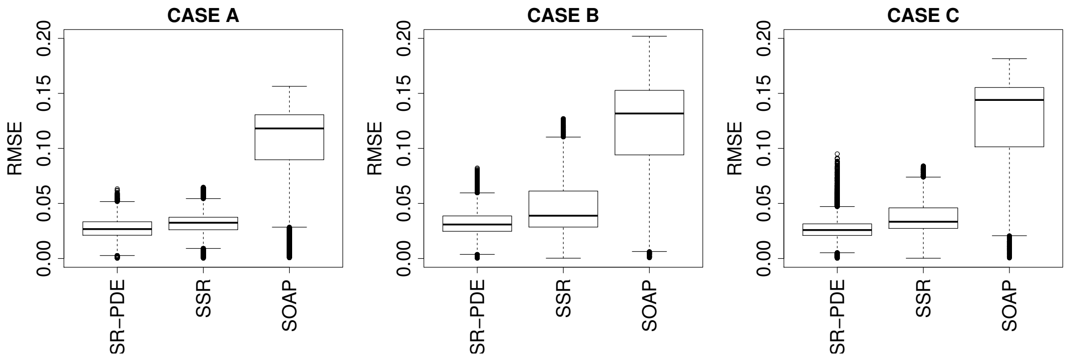

Figure 1.14: *Boxplot of RMSE (evaluated on a fine lattice of step 0.01 over the domain Ω) for SR-PDE, SSR and SOAP estimators, in case studies A, B and C (left, central and right panel, respectively).*

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 in Figures 1.11-1.13.

1.7 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 Figure 1.6, 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 $\tilde{J}(f)$ defined in (1.8). As mentioned in the Section 1.1 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 1.7 and described in equation (1.42) (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

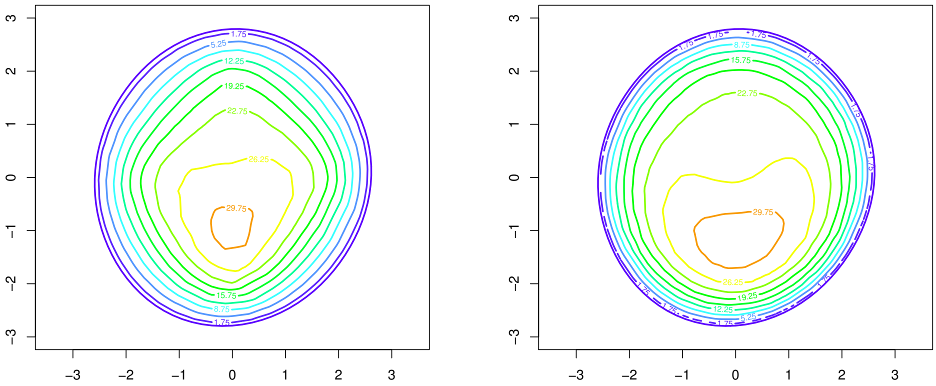

Figure 1.15: *Estimate of the blood-flow velocity field in the carotid section with standard SSR (left) and SR-PDE (right).*

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 1.7, 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 1.8.

Figure 1.15 displays the velocity field estimated using standard SSR (Left) and SR-PDE (Right) smoothing. The standard SSR estimate is obtained making collapse each beam in its central point, assuming that the ECD data are pointwise data and minimizing the functional $J(f)$ penalized with the Laplace operator. The surface estimator is then discretized by means of linear Finite Elements defined on the same mesh used in the SR-PDE estimate. A visual comparison of the surface estimate obtained with the two methods immediately highlights the advantages of the proposed technique. The standard SSR estimate is, in fact, strongly influenced by the cross-shaped pattern of the observations and displays strongly rhomboidal isolines, which are certainly non-physiological; the penalization of a measure of the local curvature of the field over-smooths and flattens the field toward a plane in those regions of the domain where no observations are available. The SR-PDE, instead, 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.

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 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 hemodynamics 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., [35]. 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.

CHAPTER 2

Mixed Finite Elements for spatial regression with PDE penalization

2.1 Introduction

In this chapter we study the properties of Spatial Regression with PDE penalization (SR-PDE) and its mixed Finite Element discretization. We recall that SR-PDE estimates bidimensional or three dimensional fields minimizing a least square functional penalized with the L^2-norm over the domain of interest of the misfit of a second order PDE, $Lf = u$, modeling the phenomenon under study. All the parameters appearing in the operator L and the boundary conditions are known while the forcing term in the PDE is not completely determined. This approach is similar to the one used in control theory when a distributed control is considered, see for example [23]. The main difference from classical results in control theory is that the observations are pointwise and affected by noise. For this reason it is necessary to require higher regularity to the field to ensure that the penalized least square functional is well defined.

The penalized least square functional has a unique minimum in the Sobolev space H^2 and the minimum is the solution of a fourth order problem. In order to prove the existence and the uniqueness of the estimator we resort to a mixed approach for fourth order problems, since the penalized error functional is not necessarily convex in H^2. Accordingly, a mixed equal order Finite Element method, similar to classical mixed methods described for example in [6], is used for discretizing the estimation problem. Other classical conforming and nonconforming methods (see [6] and references therein) or more recent discontinuous Galerkin methods (see, e.g., [3, 16, 34]) can be used for the discretization of the fourth order problem. However, in the specific case here considered the mixed Finite Element method is a convenient choice since the problem in exam can be written as a system of second order PDEs. Moreover the mixed approach provides also a good approximation of second order derivatives of the field that can be useful in order to compute physical quantities of interest.

The proposed mixed equal order Finite Elements discretization is known to have sub-optimal convergence rate when applied to fourth order problems with arbitrary boundary conditions and, in particular, the first order approximation might not converge to the exact solution (see, e.g., [4, 6]). However we are able to prove the optimal convergence of the proposed discretization method for the specific set of boundary conditions that are naturally associated to the smoothing problem, whenever the true underlying field satisfies exactly those conditions. The theoretical results are confirmed by numerical experiments.

The inspected convergence concerns the study of the bias of the estimator, while the study of the variance of the estimator and the convergence when the number of observations goes to infinity will be the subject of a future work. These topics are studied in the classical setting of smoothing splines (see, e.g., [7]), thin-plate splines or multidimensional smoothing splines (see, e.g., [8, 9, 18] and references therein) but they cannot be directly generalized to SR-PDE models.

The chapter is organized as follows. Section 2.2 recalls SR-PDE model used for pointwise observations. Section 2.3 proves the well-posedness of the estimation problem and Section 2.4 obtains a bound for the bias of the estimator. Section 2.5 describes the mixed Finite Element method used for the discretization of the estimation problem and proves the well-posedness of the discrete problem. Section 2.6 proves the convergence of the proposed mixed Finite Element method and provides a bound for the bias of the Finite Element estimator. Section 2.7 presents the numerical experiments supporting the theoretical results. Section 2.8 extends the models to the case of areal data and presents the asymptotic results in this setting.

2.2 Surface estimator for pointwise data

We generalize to dimension $d \leq 3$ the SR-PDE models presented in Chapter 1. Similarly to Section 1.2, we introduce the model for pointwise data in $\mathbb{R}^d$ with $d \leq 3$. Consider a bounded, regular, open domain $\Omega \subset \mathbb{R}^d$ with $d \leq 3$, whose boundary $\partial\Omega$ is a curve of class C^2, and a regular function $f_0 : \Omega \to \mathbb{R}$ to be estimated from noisy observations. Let z_i, for $i = 1, \dots, n$, be n observations that represent noisy evaluations of the field f_0 at points $\mathbf{p}_i \in \Omega$. The error model that we consider for the observations is a classical additive model:

$$z_i = f_0(\mathbf{p}_i) + \epsilon_i \tag{2.1}$$

where ϵ_i, $i = 1, \dots, n$, are independent errors with zero mean and constant variance σ^2.

We suppose to have, in addition to the observations z_i, a physical knowledge of the phenomenon under study and that this prior knowledge can be described by means of a differential operator. Specifically, we can formalize this as a PDE that f_0 satisfies:

$$\begin{cases} Lf_0 = \tilde{u} & in\ \Omega \\ \mathcal{B}_c f_0 = h & on\ \partial\Omega \end{cases} \tag{2.2}$$

where the operator L and the boundary conditions are completely determined and fixed, while the forcing term $\tilde{u} = u + g_0 \in L^2(\Omega)$ is composed by a known and fixed part

u and an unknown term, called g_0, that will be estimated from data. The parameters of the PDE and the boundary conditions could be as well considered partly unknown and estimated from data, but in this thesis we assume them to be known and fixed. We focus on second order elliptic operators, in particular L is a diffusion-transport-reaction operator

$$Lf_0 = -\mathrm{div}(\mathbf{K}\nabla f_0) + \mathbf{b}\cdot\nabla f_0 + cf_0 \tag{2.3}$$

with smooth and bounded parameters. The matrix $\mathbf{K} \in \mathbb{R}^{d\times d}$ is a symmetric and positive definite diffusion tensor, $\mathbf{b} \in \mathbb{R}^d$ is the transport vector and $c \geq 0$ is the reaction term. These parameters can be spatially varying in Ω; i.e., $\mathbf{K} = \mathbf{K}(\mathbf{x})$, $\mathbf{b} = \mathbf{b}(\mathbf{x})$ and $c = c(\mathbf{x})$, with $\mathbf{x} \in \Omega$. The boundary conditions of the PDE are homogeneous or non-homogeneous Dirichlet, Neumann, Robin (or mixed) conditions. All the admissible boundary conditions are summarized in

$$\mathcal{B}_c f_0 = \begin{cases} f_0 & \text{on } \Gamma_D \\ \mathbf{K}\nabla f_0\cdot\boldsymbol{\nu} & \text{on } \Gamma_N \\ \mathbf{K}\nabla f_0\cdot\boldsymbol{\nu} + \gamma f_0 & \text{on } \Gamma_R \end{cases} \qquad h = \begin{cases} h_D & \text{on } \Gamma_D \\ h_N & \text{on } \Gamma_N \\ h_R & \text{on } \Gamma_R \end{cases} \tag{2.4}$$

where $\boldsymbol{\nu}$ is the outward unit normal vector to $\partial\Omega$, $\gamma \in \mathbb{R}$ is a positive constant and $\partial\Omega = \bar{\Gamma}_D \cup \bar{\Gamma}_N \cup \bar{\Gamma}_R$, with $\Gamma_D, \Gamma_N, \Gamma_R$ not overlapping.
In what follows, we make the following assumption.

Assumption 1. $\Gamma_D \neq \emptyset$*, so that a Poincaré inequality holds, i.e.,*

$$\|v\|_{L^2(\Omega)} \leq C_P \|\nabla v\|_{L^2(\Omega)}. \tag{2.5}$$

Similarly to Chapter 1, in order to estimate the field f_0 we minimize the penalized sum-of-square-error functional

$$J(f) = \frac{1}{n}\sum_{i=1}^{n}(f(\mathbf{p}_i) - z_i)^2 + \lambda\int_\Omega (Lf - u)^2 \tag{2.6}$$

over the set of functions $V = \{v \in H^2(\Omega) : \mathcal{B}_c v = h\}$ where $H^2(\Omega)$ is the Sobolev space of functions in $L^2(\Omega)$ with first and second derivatives in $L^2(\Omega)$; notice that the boundary conditions (2.4) are imposed directly in V. Even if in this case we are considering fixed and deterministic boundary conditions, when data on the boundary are available it is possible to include the uncertainty on the boundary conditions in the model by means of a dedicated regularizing term in the least square functional.
Notice that the functional J is slightly different from the functional used in Chapter 1. Specifically, the least square term in this case is divided by n, for convenience in the numerical proofs. This corresponds to multiplying n to the roughness parameter λ in the functional (1.2).
The functional J is still composed by a data fitting criterion, consisting in classical least square errors, and a model fitting criterion, formalized as a roughness term that penalizes the misfit of the PDE governing the phenomenon. Notice that by minimizing the misfit of the PDE $Lf_0 - u$, where u is the known part of the forcing term, we are actually minimizing the contribution of the unknown forcing term g_0. The contribution

of the data fitting criterion and of the model fitting criterion is tuned by means of the parameter λ. A large literature is devoted to the optimal choice of this parameter; see, e.g., [20, 30] and references therein; classical methods are for example the Akaike's Information Criterion (AIC), the Bayesian Information Criterion (BIC) and the Generalized Cross-Validation (GCV) criterion.

The functional $J(f)$ is well defined if $f \in H^2(\Omega)$ thanks to the embedding $H^2(\Omega) \subset C(\bar{\Omega})$ if $\Omega \subset \mathbb{R}^d$ with $d \leq 3$. For data in $\mathbb{R}^d$ with $d > 3$ one has to require more regularity in order to obtain $f \in C(\Omega)$; in particular one needs $f \in H^s(\Omega)$ with $s > d/2$; see, e.g., [12].
The estimation problem is formulated as follows.

Problem 3. *Find $\hat{f} \in V$ such that*

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

As it will be shown in the next section, this problem is well posed if we assume some regularity on the parameters of the PDE and on the domain Ω. In particular, in the case $d \leq 3$, we make the following assumption.

Assumption 2. *The parameters of the PDE are such that $\forall \tilde{u} \in L^2(\Omega)$ there exists a unique solution f_0 of the PDE* (2.2)*, which moreover satisfies $f_0 \in H^2(\Omega)$.*

The Lax-Milgram theorem guarantees the existence and the uniqueness of the solution of the PDE (2.2) in $H^1(\Omega)$ when the parameters of the PDE $\mathbf{K}$, $\mathbf{b}$ and c satisfy some classical requests, for example $\mathbf{K}_{ij},\ \mathbf{b}_j,\ c \in L^\infty(\Omega)$, $\mathbf{K}$ is symmetric and uniformly elliptic, i.e., $\boldsymbol{\xi}^T\mathbf{K}(\mathbf{x})\boldsymbol{\xi} \geq \alpha_K\ \forall \mathbf{x} \in \Omega$ and $\forall \boldsymbol{\xi} \in \mathbb{R}^d$, $\mathbf{b} \cdot \boldsymbol{\nu} \geq 0$ on $\Gamma_N \cup \Gamma_R$, $-1/2\mathrm{div}(\mathbf{b}(\mathbf{x})) + c(\mathbf{x}) \geq -\alpha_K/C_P$, where α_K is the ellipticity constant and C_P is the Poincaré constant, $\gamma \in L^\infty(\partial\Omega)$, $\gamma \geq 0$ and $h_D \in H^{1/2}(\partial\Omega)$, $h_N \in H^{-1/2}(\partial\Omega)$, $h_R \in H^{-1/2}(\partial\Omega)$.
To guarantee that the solution of the PDE is in $H^2(\Omega)$ we need to make further assumptions on the parameters of the PDE and on the boundary conditions requiring extra regularity: $\mathbf{K}_{ij}$ is Lipschitz continuous, $h_D \in H^{3/2}(\partial\Omega)$, $h_N \in H^{1/2}(\partial\Omega)$, $h_R \in H^{1/2}(\partial\Omega)$.
If the boundary conditions imposed are mixed, they have to satisfy some joint conditions in order not to reduce the regularity of the solution; see [12] for more details.

2.3 Well posedness analysis

To analyze the well-posedness of Problem 3 we introduce a new quantity $g \in \mathcal{G} = L^2(\Omega)$ that represents the misfit of the PDE in the penalizing term. This new quantity, $g \in \mathcal{G}$, is defined as $g = Lf - u$, where L is the second order elliptic operator (2.3), and is the classical *control* term in PDE optimal control theory.
It is useful to introduce also the space $V_0 = \{v \in V\ :\ \mathcal{B}_c v = 0\}$, which represents the space of functions in V with homogeneous boundary conditions, and the operator $B : L^2(\Omega) \to V_0$ such that $B\tilde{u}$ is the unique solution of the PDE (2.2) with forcing term $\tilde{u}$ and homogeneous boundary conditions, i.e., $L(B\tilde{u}) = \tilde{u}$ in Ω and $\mathcal{B}_c(B\tilde{u}) = 0$ on $\partial\Omega$. Under Assumptions 1 and 2, thanks to the well-posedness and the H^2-regularity of the PDE (2.2), the operator B is an isomorphism between the spaces L^2 and V_0 and the

H^2-norm of Bu is equivalent to the L^2-norm of u, i.e., there exist two positive constants C_1 and C_2 such that

$$C_1 \|u\|_{L^2(\Omega)} \leq \|Bu\|_{H^2(\Omega)} \leq C_2 \|u\|_{L^2(\Omega)} . \tag{2.7}$$

The solution of the PDE (2.2) can thus be written as $f = f_b + B\tilde{u}$, where f_b is the solution of the PDE with homogeneous forcing term and non-homogeneous boundary conditions.
Existence and uniqueness of the estimator $\hat{f}$ is obtained thanks to classical results of calculus of variations. We recall here the result stated, e.g., in [23].

Theorem 1. *If the functional $J(g)$ has the form*

$$J(g) = \mathcal{A}(g, g) + \mathcal{L}g + c \tag{2.8}$$

where $\mathcal{A} : \mathcal{G} \times \mathcal{G} \to \mathbb{R}$ is a continuous, coercive and symmetric bilinear form in $\mathcal{G}$, $\mathcal{L} : \mathcal{G} \to \mathbb{R}$ is a linear operator, c is a constant and $\mathcal{G}$ is a Hilbert space, then $\exists!\ \hat{g} \in \mathcal{G}$ such that $J(\hat{g}) = \inf_{\mathcal{G}} J(g)$.
Moreover $\hat{g}$ satisfies the following Euler-Lagrange equation:

$$(J'(\hat{g}), \varphi) = 2\mathcal{A}(\hat{g}, \varphi) + \mathcal{L}\varphi = 0 \ \ \forall \varphi \in \mathcal{G}. \tag{2.9}$$

The existence and uniqueness of the estimator is stated in the following theorem.

Theorem 2. *Under Assumptions 1 and 2, the solution of Problem 3 exists and is unique.*

Proof. Thanks to the definition of g we can write f as an affine transformation of g, i.e., $f = f_b + B(u + g)$, and the functional (2.6) as

$$J_g(g) = J(f_b + B(u + g)) = \frac{1}{n} \sum_{i=1}^{n} (B(u + g)(\mathbf{p}_i) + f_b(\mathbf{p}_i) - z_i)^2 + \lambda \|g\|^2_{L^2(\Omega)} . \tag{2.10}$$

This reformulation of the functional J is very useful since we can now write J_g in the quadratic form (2.8) where

$$\begin{aligned}
\mathcal{A}(g, \varphi) =& \frac{1}{n} \sum_{i=1}^{n} Bg(\mathbf{p}_i) B\varphi(\mathbf{p}_i) + \lambda \int_\Omega g\varphi \\
\mathcal{L}\varphi =& \frac{2}{n} \sum_{i=1}^{n} B\varphi(\mathbf{p}_i)(Bu(\mathbf{p}_i) + f_b(\mathbf{p}_i) - z_i) \\
c =& \frac{1}{n} \sum_{i=1}^{n} (Bu(\mathbf{p}_i) + f_b(\mathbf{p}_i) - z_i)^2.
\end{aligned}$$

Clearly $\mathcal{A}(g, \varphi)$ is a bilinear form, since both B and the pointwise evaluation of a function are linear operators. Moreover, it is continuous in $\mathcal{G}$; indeed, thanks to the embedding $H^2(\Omega) \subset C(\bar{\Omega})$ if $\Omega \subset \mathbb{R}^d$ with $d \leq 3$ and thanks to (2.7) we have that

$$|Bg(\mathbf{p}_i)| \leq \|Bg\|_{C(\bar{\Omega})} \leq C \|Bg\|_{H^2(\Omega)} \leq C \|g\|_{L^2(\Omega)} .$$

We thus obtain that $\mathcal{A}(g, \varphi) \leq (C^2 + \lambda) \|g\|_{L^2(\Omega)} \|\varphi\|_{L^2(\Omega)}$.
Finally, the operator $\mathcal{A}(g, \varphi)$ is coercive in $L^2(\Omega)$, since

$$\mathcal{A}(g, g) = \frac{1}{n} \sum_{i=1}^{n} |Bg(\mathbf{p}_i)|^2 + \lambda \int_\Omega g^2 \geq \lambda \int_\Omega g^2 = \lambda \|g\|^2_{L^2(\Omega)} .$$

Due to the fact that the bilinear form $\mathcal{A}(\cdot,\cdot)$ is continuous and coercive in $\mathcal{G} = L^2(\Omega)$, that the operator $\mathcal{L}$ is linear and that c is a constant, Theorem 1 states the existence and the uniqueness of $\hat{g} = \operatorname{argmin}_{g\in\mathcal{G}} J_g(g)$. From the bijectivity of $B : L^2(\Omega) \to V_0$ we deduce the existence and uniqueness of $\hat{f} = f_b + B(\hat{g} + u) = \operatorname{argmin}_{f\in V} J(f)$. $\square$

The estimator $\hat{f}$ is obtained by solving:

$$\begin{cases} L\hat{f} = u + \hat{g} & in\ \Omega \\ \mathcal{B}_c\hat{f} = h & on\ \partial\Omega. \end{cases} \tag{2.11}$$

We now show that if $\hat{g}$ is smooth enough, e.g., $\hat{g} \in H^2(\Omega)$, then $\hat{g}$ solves the PDE

$$\begin{cases} L^*\hat{g} = -\frac{1}{n\lambda} \sum_{i=1}^{n} (\hat{f} - z_i)\delta_{\mathbf{p}_i} & in\ \Omega \\ \mathcal{B}_c{}^*\hat{g} = 0 & on\ \partial\Omega \end{cases} \tag{2.12}$$

where $\delta_{\mathbf{p}_i}$ is the Dirach mass located in $\mathbf{p}_i$, L^* is the adjoint operator of L

$$L^*g = -\text{div}(\mathbf{K}\nabla g) - \mathbf{b}\cdot\nabla g + (c - \text{div}(\mathbf{b}))g \tag{2.13}$$

and the “adjoint” boundary conditions are

$$\mathcal{B}_c{}^*g = \begin{cases} g & \text{on } \Gamma_D \\ \mathbf{K}\nabla g\cdot\boldsymbol{\nu} + \mathbf{b}\cdot\boldsymbol{\nu} g & \text{on } \Gamma_N \\ \mathbf{K}\nabla g\cdot\boldsymbol{\nu} + (\mathbf{b}\cdot\boldsymbol{\nu} + \gamma)g & \text{on } \Gamma_R. \end{cases} \tag{2.14}$$

Indeed, if we associate to any $\varphi \in \mathcal{G}$ the function $v \in V_0$ such that $v = B\varphi$ (or equivalently $Lv = \varphi$ with homogeneous boundary conditions), we have that

$$\begin{aligned} \frac{1}{2}\left(J'_g(\hat{g}), \varphi\right) &= \int_\Omega \frac{1}{n}\sum_{i=1}^{n}\left(\hat{f} - z_i\right) v\delta_{\mathbf{p}_i} + \lambda\int_\Omega \hat{g}Lv \\ &= \int_\Omega \frac{1}{n}\sum_{i=1}^{n}\left(\hat{f} - z_i\right) v\delta_{\mathbf{p}_i} + \lambda\int_\Omega L^*\hat{g}v + \int_{\partial\Omega}\hat{g}\mathcal{B}_c v + \int_{\partial\Omega}\mathcal{B}_c{}^*\hat{g}v \\ &= \int_\Omega \left[\frac{1}{n}\sum_{i=1}^{n}\left(\hat{f} - z_i\right)\delta_{\mathbf{p}_i} + \lambda L^*\hat{g}\right] B\varphi + \int_{\partial\Omega}\mathcal{B}_c{}^*\hat{g}B\varphi = 0\ \ \forall\varphi\in\mathcal{G}. \end{aligned}$$

Since B is a bijection between $L^2(\Omega)$ and V_0 the latter is equivalent to

$$\int_\Omega \left[\frac{1}{n}\sum_{i=1}^{n}\left(\hat{f} - z_i\right)\delta_{\mathbf{p}_i} + \lambda L^*\hat{g}\right] v + \int_{\partial\Omega}\mathcal{B}_c{}^*\hat{g}v = 0\ \ \forall v \in V_0. \tag{2.15}$$

Choosing $v \in C^{\infty}$ with compact support in Ω, equation (2.15) implies that $\hat{g}$ is the solution in the sense of distributions of

$$L^* \hat{g} = -\frac{1}{n\lambda} \sum_{i=1}^{n} (\hat{f} - z_i) \delta_{\mathbf{p}_i}.$$

Choosing $v \in C^{\infty}(\bar{\Omega})$ with support not including any of the location points $\mathbf{p}_i$, we obtain from equation (2.15) the boundary conditions for $\hat{g}$ that are ${\mathcal{B}_c}^* \hat{g} = 0$ where ${\mathcal{B}_c}^*$ is defined in equation (2.14).
In this case, the estimator $\hat{f}$ is obtained by solving the coupled system of PDEs

$$\begin{cases} L\hat{f} = u + \hat{g} & in\ \Omega \\ \mathcal{B}_c \hat{f} = h & on\ \partial\Omega \end{cases} \qquad \begin{cases} L^* \hat{g} = -\frac{1}{n\lambda} \sum_{i=1}^{n} (\hat{f} - z_i) \delta_{\mathbf{p}_i} & in\ \Omega \\ {\mathcal{B}_c}^* \hat{g} = 0 & on\ \partial\Omega. \end{cases} \tag{2.16}$$

2.4 Bias of the estimator

The penalty term in the functional $J(f)$ induces a bias in the estimator $\hat{f}$ unless the unknown part of the forcing term $g_0 = 0$ and the true underlying field f_0 satisfies exactly the penalized PDE $Lf_0 = u$; we want now to quantify this bias.
The estimator $\hat{f}$ is obtained as the unique minimum of the functional $J(f)$, solving the Euler-Lagrange equation (2.9). Thanks to the linearity of equation (2.9), we can write

$$\begin{aligned} \hat{f} = & \operatorname*{argmin}_{f \in V} \left[\frac{1}{n} \sum_{i=1}^{n} (f(\mathbf{p}_i) - f_0(\mathbf{p}_i))^2 + \lambda \int_{\Omega} (Lf - u)^2 \right] \\ & + \operatorname*{argmin}_{w \in V_0} \left[\frac{1}{n} \sum_{i=1}^{n} (w(\mathbf{p}_i) - \epsilon_i)^2 + \lambda \int_{\Omega} (Lw)^2 \right] \end{aligned}$$

where $f_0(\mathbf{p}_i)$ is the mean value of the observation z_i located in $\mathbf{p}_i$, i.e., $\mathbb{E}[z_i] = f_0(\mathbf{p}_i)$. The first term is the deterministic part of $\hat{f}$, while the second term

$$\hat{w} = \operatorname*{argmin}_{w \in V_0} \left[\frac{1}{n} \sum_{i=1}^{n} (w(\mathbf{p}_i) - \epsilon_i)^2 + \lambda \int_{\Omega} (Lw)^2 \right] \tag{2.17}$$

is related to the observation noise: $\hat{w}$ is in fact the minimizer of the functional when data are pure noise and the penalized PDE is homogeneous (both the forcing term and the boundary conditions are homogeneous). Notice that $\hat{w}$ is a linear function of the errors ϵ_i and for this reason it has zero mean. Indeed, $\hat{w}$ is obtained as the solution of the PDE

$$\begin{cases} L\hat{w} = \hat{g}_w & in\ \Omega \\ {\mathcal{B}_c}^* \hat{w} = 0 & on\ \partial\Omega \end{cases} \tag{2.18}$$

where $\hat{g}_w$ satisfies

$$\begin{cases} L^* \hat{g}_w = -\frac{1}{n\lambda} \sum_{i=1}^{n} (\hat{w} - \epsilon_i) \delta_{\mathbf{p}_i} & in\ \Omega \\ {\mathcal{B}_c}^* \hat{g}_w = 0 & on\ \partial\Omega. \end{cases} \tag{2.19}$$

Moreover, thanks to the PDE (2.18) we know that $\mathbb{E}[\hat{w}] = B\mathbb{E}[\hat{g}_w]$, while from the PDE (2.19) we obtain that

$$L^*\mathbb{E}[\hat{g}_w] + \frac{1}{n\lambda}\sum_{i=1}^{n} B\mathbb{E}[\hat{g}_w(\mathbf{p}_i)]\delta_{\mathbf{p}_i} = \frac{1}{n\lambda}\sum_{i=1}^{n}\mathbb{E}[\epsilon_i]\delta_{\mathbf{p}_i} = 0.$$

Finally, since L^*, B and the evaluation in a point are linear operators, we have that both $\hat{g}_w$ and $\hat{w}$ have zero mean. It follows then that the mean value of the estimator $\mathbb{E}[\hat{f}]$ is the minimizer of the functional when the observations are without noise and the PDE is not homogeneous (both the forcing term and the boundary conditions are not homogeneous), i.e.,

$$\mathbb{E}[\hat{f}] = \underset{f\in V}{\operatorname{argmin}}\left[\frac{1}{n}\sum_{i=1}^{n}\left(f(\mathbf{p}_i) - f_0(\mathbf{p}_i)\right)^2 + \lambda\int_\Omega (Lf-u)^2\right]. \tag{2.20}$$

Notice that the sum of the Euler-Lagrange equations associated to the functionals (2.17) and (2.20) corresponds to the Euler-Lagrange equation (2.9) associated to the functional (2.6).

Since $\mathbb{E}[\hat{f}]$ is related to the bias induced by the penalizing term, we are interested in studying the error term $\mathbb{E}[\hat{f} - f_0]$; in particular it is natural to study it in the norm induced by the functional $J(f)$, i.e.,

$$\|f\|_J^2 = \frac{1}{n}\sum_{i=1}^{n} f(\mathbf{p}_i)^2 + \lambda\int_\Omega (Lf)^2. \tag{2.21}$$

Lemma 1. *The norm* (2.21) *of the bias of* $\hat{f}$ *is bounded by*

$$\left\|\mathbb{E}[\hat{f} - f_0]\right\|_J^2 \leq 4\lambda \left\|Lf_0 - u\right\|_{L^2(\Omega)}^2. \tag{2.22}$$

Proof. In order to obtain the inequality (2.22) we can use the optimality (2.20) of $\mathbb{E}[\hat{f}]$ in the minimization of the functional with respect to any other function in V. We have in fact that

$$\begin{aligned}
&\left\|\mathbb{E}[\hat{f}] - f_0\right\|_J^2 = \frac{1}{n}\sum_{i=1}^{n}(\mathbb{E}[\hat{f}](\mathbf{p}_i) - f_0(\mathbf{p}_i))^2 + \lambda\left\|L(\mathbb{E}[\hat{f}] - f_0)\right\|_{L^2(\Omega)}^2 \\
&\leq \frac{1}{n}\sum_{i=1}^{n}(\mathbb{E}[\hat{f}](\mathbf{p}_i) - f_0(\mathbf{p}_i))^2 + 2\lambda\left\|L\mathbb{E}[\hat{f}] - u\right\|_{L^2(\Omega)}^2 + 2\lambda\left\|Lf_0 - u\right\|_{L^2(\Omega)}^2 \\
&\leq 2\left[\frac{1}{n}\sum_{i=1}^{n}(\mathbb{E}[\hat{f}](\mathbf{p}_i) - f_0(\mathbf{p}_i))^2 + \lambda\left\|L\mathbb{E}[\hat{f}] - u\right\|_{L^2(\Omega)}^2\right] + 2\lambda\left\|Lf_0 - u\right\|_{L^2(\Omega)}^2 \\
&\leq 2\lambda\left\|Lf_0 - u\right\|_{L^2(\Omega)}^2 + 2\lambda\left\|Lf_0 - u\right\|_{L^2(\Omega)}^2.
\end{aligned}$$

□

This result means that the estimator is asymptotically unbiased in the norm $\|\cdot\|_J$ either if $\|Lf_0 - u\|_{L^2(\Omega)} = 0$ or if $\lambda \to 0$ for $n \to +\infty$. The condition $\|Lf_0 - u\|_{L^2(\Omega)} = 0$ means that the real field f_0 is in the kernel of the penalty term, while the condition $\lambda \to 0$ for $n \to +\infty$ means that the more observations we have, the less we penalize the PDE misfit.

2.5 Finite Element estimator

The estimation problem presented in Section 2.2 is infinite dimensional and cannot be solved analytically. To reduce this infinite dimensional problem to a finite dimensional one we approximate the PDE system (2.16) with the Finite Element method; this method has already been used in this framework for example in [11,31,32]. The Finite Element approximation of the system (2.16) can be regarded as a naive mixed Finite Element method for the discretization of Problem 3. More complex methods for the discretization of fourth order problems could be used: in [6], for example, some conforming and nonconforming methods for the discretization of fourth order problems are introduced, while in [3, 16, 34] more recent discontinuous Galerkin methods are described.

Let $\mathcal{T}_h$ be a regular and quasi-uniform triangulation of the domain, that for convenience we assume here to be polygonal and convex, and $h = \max_{K\in\mathcal{T}_h} \mathrm{diam}(K)$ be the characteristic mesh size (see, e.g., [2]). Notice that the mesh $\mathcal{T}_h$ can be defined independently of the location of the observations $\mathbf{p}_1, \cdots, \mathbf{p}_n$. We consider the space V_h^r of piecewise continuous polynomial functions of degree $r \geq 1$ on the triangulation

$$V_h^r = \left\{ v \in C^0(\bar{\Omega}) : v|_K \in \mathbb{P}^r(K)\ \forall K \in \mathcal{T}_h \right\}$$

and $V^r_{h,\Gamma_D} = V_h^r \cap H^1_{\Gamma_D}(\Omega)$ where $H^1_{\Gamma_D} = \{v \in H^1(\Omega) :\ v|_{\Gamma_D} = 0\}$.
In order to discretize the PDE system (2.16) we define the bilinear forms

$$r(g, v) = \int_\Omega gv, \quad l(f, \psi) = \frac{1}{n}\sum_{i=1}^n f(\mathbf{p}_i)\psi(\mathbf{p}_i),$$

$$a(f, \psi) = \int_\Omega (\mathbf{K}\nabla f \cdot \nabla\psi + \mathbf{b}\cdot\nabla f\psi + cf\psi) + \int_{\Gamma_R} \gamma f\psi, \tag{2.23}$$

the latter being the bilinear form associated to the operator L; we also introduce the linear operator $F(\psi) = \int_\Omega u\psi + \int_{\Gamma_N} h_N\psi + \int_{\Gamma_R} h_R\psi$.
Let now $f_{D,h} \in V_h^r$ be a lifting of the non-homogeneous Dirichlet conditions, i.e, $f_{D,h}|_{\Gamma_D} = h_{D,h}$, where $h_{D,h}$ is the interpolant of h_D in the space of piecewise continuous polynomial functions of degree r on the Dirichlet boundary Γ_D. The Finite Element approximation of the system (2.16) becomes

$$\begin{cases} \frac{1}{\lambda} l(\hat{f}_h, \psi_h) + a(\psi_h, \hat{g}_h) = \frac{1}{n\lambda}\sum_{i=1}^n z_i\psi_h(\mathbf{p}_i) & \forall \psi_h \in V^r_{h,\Gamma_D} \\ a(\hat{f}_h, v_h) - r(\hat{g}_h, v_h) = F(v_h) & \forall v_h \in V^r_{h,\Gamma_D} \end{cases} \tag{2.24}$$

with $(\hat{f}_h - f_{D,h}, \hat{g}_h) \in V^r_{h,\Gamma_D} \times V^r_{h,\Gamma_D}$.

In this section and in the following one, we need a slightly stronger regularity assumption on the PDE (2.2), in particular we need its solution to be in a Sobolev space $W^{2,p}$, where $W^{s,p}(\Omega)$ is the space of functions in $L^p(\Omega)$ with derivatives up to order s in $L^p(\Omega)$.

Assumption 3. *The parameters of the PDE are such that* $\forall \tilde{u} \in L^p(\Omega)$ *there exists a unique solution* $f_0 \in W^{2,p}(\Omega)$, *for some* $p > d$.

Lemma 2. *Under Assumption 3, there exists* $h_0 > 0$ *s.t.* $\forall h \leq h_0$, *Problem* (2.24) *has a unique solution.*

Proof. The proof mimics the strategy used to prove the existence and the uniqueness of the estimator at the continuos level in Theorem 2.
Let $B : L^2(\Omega) \to V_0$ be the operator defined in Section 2.3 such that $\psi = B\varphi_h$ is the solution of

$$a(\psi, v) = \int_\Omega \varphi_h v \ \ \forall v \in H^1_{\Gamma_D}.$$

We define the operator B_h as the discretization of the operator B, i.e., $\psi_h = B_h\varphi_h \in V^r_{h,\Gamma_D}$ is the solution of

$$a(\psi_h, v_h) = \int_\Omega \varphi_h v_h \ \ \forall v_h \in V^r_{h,\Gamma_D}.$$

It is easy to show that the operator B_h is stable in the L^∞-norm, i.e., $\|\psi_h\|_{L^\infty(\Omega)} \leq C\|\varphi_h\|_{L^2(\Omega)}$. We have in fact that

$$\|\psi_h\|_{L^\infty(\Omega)} \leq \|\psi - \psi_h\|_{L^\infty(\Omega)} + \|\psi\|_{L^\infty(\Omega)}.$$

Thanks to the H^2-elliptic regularity of the PDE (2.2) (see Assumption 2) we have that

$$\|\psi\|_{L^\infty(\Omega)} \leq C\|\psi\|_{H^2(\Omega)} \leq C\|\varphi_h\|_{L^2(\Omega)}$$

while thanks to Assumption 3 and the Sobolev inequality (see, e.g., [2])

$$\|w\|_{L^\infty(\Omega)} \leq C\|w\|_{W^{1,p}(\Omega)} \quad \forall w \in W^{1,p}(\Omega) \quad \forall p > d \tag{2.25}$$

where $W^{1,p}(\Omega)$ is the space of functions in $L^p(\Omega)$ with first derivatives in $L^p(\Omega)$, we obtain the bound for the error term in the L^∞-norm

$$\begin{aligned}\|\psi - \psi_h\|_{L^\infty(\Omega)} &\leq C\|\psi - \psi_h\|_{W^{1,p}(\Omega)} \leq C \inf_{v_h \in V^r_{h,\Gamma_D}} \|\psi - v_h\|_{W^{1,p}(\Omega)} \\ &\leq Ch\,|\psi|_{W^{2,p}(\Omega)} \leq Ch\,\|\varphi_h\|_{L^p(\Omega)} \leq Ch^{1+\min\left\{0,\frac{d}{p}-\frac{d}{2}\right\}}\|\varphi_h\|_{L^2(\Omega)}.\end{aligned}$$

In the last step we have used an inverse inequality; see, e.g., [10], and taking $p = 2d/(d-2)$, which is larger than d for $d \leq 3$, we conclude

$$\|\psi - \psi_h\|_{L^\infty(\Omega)} \leq C\|\varphi_h\|_{L^2(\Omega)}.$$

We define now the operators $\mathcal{A}_h$ and $\mathcal{L}_h$ as the discretization of the operators $\mathcal{A}$ and $\mathcal{L}$ defined in Section 2.3:

$$\begin{aligned}\mathcal{A}_h(g_h, \varphi_h) &= \frac{1}{n}\sum_{i=1}^n B_h g_h(\mathbf{p}_i) B_h\varphi_h(\mathbf{p}_i) + \lambda\int_\Omega g_h\varphi_h \\ \mathcal{L}_h\varphi_h &= \frac{2}{n}\sum_{i=1}^n B_h\varphi_h(\mathbf{p}_i)(B_h u(\mathbf{p}_i) + f_{b,h}(\mathbf{p}_i) - z_i)\end{aligned}$$

where $f_{b,h}$ is the discretization of f_b. The operator $\mathcal{A}_h$ is coercive in L^2, in fact

$$\mathcal{A}_h(g_h, g_h) = \frac{1}{n}\sum_{i=1}^n (B_h g_h(\mathbf{p}_i))^2 + \lambda\int_\Omega g_h^2 \geq \lambda\|g_h\|^2_{L^2(\Omega)}.$$

Thanks to the stability in the L^∞-norm of the operator B_h, both the operators $\mathcal{A}_h$ and $\mathcal{L}_h$ are continuous:

$$\begin{aligned}\mathcal{A}_h(g_h,\varphi_h) &\leq C\left\|B_h g_h\right\|_{L^\infty(\Omega)}\left\|B_h\varphi_h\right\|_{L^\infty(\Omega)}+\lambda\left\|g_h\right\|_{L^2(\Omega)}\left\|\varphi_h\right\|_{L^2(\Omega)}\\ &\leq C\left\|g_h\right\|_{L^2(\Omega)}\left\|\varphi_h\right\|_{L^2(\Omega)}\\ \mathcal{L}_h(\varphi_h) &\leq C\left\|B_h\varphi_h\right\|_{L^\infty(\Omega)}\leq C\left\|\varphi_h\right\|_{L^2(\Omega)}.\end{aligned}$$

Thanks to the fact that $\mathcal{A}_h$ is continuous and coercive in $L^2(\Omega)$ and $\mathcal{L}_h$ is continuous in $L^2(\Omega)$, the equation

$$2\mathcal{A}_h(g_h,\varphi_h)+\mathcal{L}_h(\varphi_h)=0\ \ \forall\varphi_h: B_h\varphi_h=\psi_h\in V^r_{h,\Gamma_D}$$

has a unique solution $g_h\in V^r_{h,\Gamma_D}$. This equation corresponds to the first equation of the system (2.24).
Once $\hat{g}_h$ is known, $\hat{f}_h$ is recovered uniquely from the second equation in (2.24). □

Let now $\{\psi_k\}_{k=1}^{N_h}$ be the Lagrangian basis of the space V^r_{h,Γ_D} of dimension $N_h=\dim(V^r_{h,\Gamma_D})$ and let $\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_{N_h}$ be the nodes associated to the N_h basis functions. Thanks to the Lagrangian property of the basis functions we can write a function $f\in\text{span}\{\psi_1,\ldots,\psi_{N_h}\}$ as

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

where $\mathbf{f}=(f_1,\ldots,f_{N_h})^T=(f(\boldsymbol{\xi}_1),\ldots,f(\boldsymbol{\xi}_{N_h}))^T$ and $\boldsymbol{\psi}=(\psi_1,\ldots,\psi_{N_h})^T$.
Analogously, we define the Lagrangian basis of the space $V^r_h\backslash V^r_{h,\Gamma_D}$ as $\{\psi^D_k\}_{k=1}^{N_h^D}$, where $N_h^D=\dim(V^r_h\backslash V^r_{h,\Gamma_D})$ and the nodes on the boundary Γ_D as $\boldsymbol{\xi}^D_1,\ldots,\boldsymbol{\xi}^D_{N_h^D}$. A lifting $f_{D,h}$ can be constructed in $\text{span}\{\psi^D_1,\ldots,\psi^D_{N_h^D}\}$ as $f_{D,h}=\mathbf{f}_D^T\boldsymbol{\psi}^D$ where $\mathbf{f}_D=(f_D(\boldsymbol{\xi}^D_1),\ldots,f_D(\boldsymbol{\xi}^D_{N_h^D}))^T$ and $\boldsymbol{\psi}^D=(\psi^D_1,\ldots,\psi^D_{N_h^D})^T$.
The Finite Element solution $\hat{f}_h$ of the discrete counterpart of the estimation problem can thus be written as

$$\hat{f}_h=\hat{\mathbf{f}}^T\boldsymbol{\psi}+\mathbf{f}_D^T\boldsymbol{\psi}^D$$

where $\hat{\mathbf{f}}$ is the solution of the linear system

$$\begin{bmatrix}\boldsymbol{\Psi}^T\boldsymbol{\Psi}/(n\lambda) & \mathbf{A}^T\\ \mathbf{A} & -\mathbf{R}\end{bmatrix}\begin{bmatrix}\hat{\mathbf{f}}\\ \hat{\mathbf{g}}\end{bmatrix}=\begin{bmatrix}\boldsymbol{\Psi}^T\mathbf{z}/(n\lambda)-\boldsymbol{\Psi}^T\boldsymbol{\Psi}^D\mathbf{f}_D/(n\lambda)\\ \mathbf{u}+\mathbf{h}_N+\mathbf{h}_R-\mathbf{A}^D\mathbf{f}_D\end{bmatrix}. \tag{2.26}$$

$\mathbf{R}_{jk}=\int_\Omega\psi_j\psi_k$ is the mass matrix, $\boldsymbol{\Psi}_{ij}=\psi_j(\mathbf{p}_i)$ and $\boldsymbol{\Psi}^D_{ij}=\psi^D_j(\mathbf{p}_i)$ are the matrices of pointwise evaluation of the basis functions, $\mathbf{A}_{jk}=a(\psi_k,\psi_j)$ and $\mathbf{A}^D_{jk}=a(\psi^D_k,\psi_j)$ are the matrices associated to the bilinear form $a(\cdot,\cdot)$. The vector $\mathbf{z}=(z_1,\ldots,z_n)$ contains the observed data while the vectors $\mathbf{u}_j=\int_\Omega u\psi_j$, $(\mathbf{h}_N)_j=\int_{\Gamma_N}h_N\psi_j$ and $(\mathbf{h}_R)_j=\int_{\Gamma_R}h_R\psi_j$ are related to the forcing term and the non homogeneous boundary conditions.

2.6 Bias of the Finite Element estimator

The Finite Element estimator $\hat{f}_h$ can be split, as its continuous counterpart $\hat{f}$, in two different terms $\mathbb{E}[\hat{f}_h]$ and $\hat{w}_h$ that are respectively the Finite Element approximation of $\mathbb{E}[\hat{f}]$ and $\hat{w}$. Reasoning as for the continuous problem, we can easily show that $\mathbb{E}[\hat{w}_h] = 0$. Neglecting this zero mean term, we now aim at studying the bias of the Finite Element estimator, $\mathbb{E}[\hat{f}_h - f_0]$, in the norm

$$\left|\left|\left|\mathbb{E}[\hat{f}_h - f_0]\right|\right|\right|^2 = \left\|\mathbb{E}[\hat{f}_h - f_0]\right\|_n^2 + \lambda \left[\left\|\mathbb{E}[\hat{f}_h] - f_0\right\|_{H^1(\Omega)}^2 + \|\mathbb{E}[\hat{g}_h] - g_0\|_{L^2(\Omega)}^2\right] \tag{2.27}$$

where the norm $\|\cdot\|_n$ is the norm induced by the bilinear operator $l(\cdot,\cdot)$, defined as

$$\left\|\mathbb{E}[\hat{f}_h - f_0]\right\|_n = \frac{1}{n}\sum_{i=1}^{n}(\mathbb{E}[\hat{f}_h](\mathbf{p}_i) - f_0(\mathbf{p}_i))^2.$$

Notice that the norm $|||\cdot|||$ contains both the norm $\|\cdot\|_J$ and the H^1-norm of $\hat{f}_h$. We need in fact also an explicit control on the H^1-norm of $\hat{f}_h$ to study the convergence properties of the mixed Finite Element solution of the system (2.24).

Remark 3. *One might be tempted to compare $\mathbb{E}[\hat{f}_h]$ to its continuous counterpart $\mathbb{E}[\hat{f}]$. However, due to the presence of $\delta_{\mathbf{p}_i}$ in the forcing term of the dual equation in* (2.16)*, $\mathbb{E}[\hat{f}]$ is not smooth in general. For this reason, in the error analysis proposed in this section, we directly compare $\mathbb{E}[\hat{f}_h]$ with the true underlying field f_0, which is assumed to be sufficiently smooth.*

The convergence of the bias term is studied when $h \to 0$, fixing the number n of observations and the penalty parameter λ. Since we are considering n and λ fixed we expect to obtain an error bound that contains a term going to zero as $h \to 0$ and a term that represents the bias induced by the roughness penalty, similarly to the continuous setting in Lemma 1.

Thanks to the introduction of the adjoint variable $\hat{g}$, which represents the misfit of the PDE, the estimation Problem 3 can be reformulated as a constrained problem that is more convenient for the study of the convergence of the Finite Element estimator.

Problem 4. *Find $\hat{f} \in V$, $\hat{g} \in \mathcal{G}$ such that*

$$(\hat{f}, \hat{g}) = \underset{(f,g)\in\mathcal{W}}{\operatorname{argmin}} \frac{1}{n}\sum_{i=1}^{n}(f(\mathbf{p}_i) - z_i)^2 + \lambda \int g^2$$

where $\mathcal{W}$ is the constrained space

$$\mathcal{W} = \{(f,g) \in V \times \mathcal{G} : \; Lf - u = g\}. \tag{2.28}$$

The constrained space $\mathcal{W}$, can be discretized as

$$\mathcal{W}_h = \{(f_h, g_h) \in V_h^r \times V_{h,\Gamma_D}^r : f_h|_{\Gamma_D} = h_{D,h} \text{ and } a(f_h, v_h) - r(g_h, v_h) = F(v_h), \; \forall v_h \in V_{h,\Gamma_D}^r\},$$

where $a(\cdot,\cdot)$, $r(\cdot,\cdot)$, $F(\cdot)$ and $h_{D,h}$ are defined in Section 2.5. The expected value of the Finite Element estimator $(\mathbb{E}[\hat{f}_h], \mathbb{E}[\hat{g}_h])$ is thus the solution of the equation

$$l(\mathbb{E}[\hat{f}_h], \psi_h) + \lambda a(\psi_h, \mathbb{E}[\hat{g}_h]) = l(f_0, \psi_h) \;\; \forall \psi_h \in V^r_{h,\Gamma_D} \tag{2.29}$$

in the constrained space $\mathcal{W}_h$.
The bound for the bias of the Finite Element estimator $\mathbb{E}[\hat{f}_h - f_0]$ is obtained thanks to the following Lemma and Theorem.

Lemma 3. *Let $g_0 = Lf_0 - u$. The bias of the Finite Element estimator $(\mathbb{E}[\hat{f}_h], \mathbb{E}[\hat{g}_h]) \in \mathcal{W}_h$ satisfies the inequality*

$$\begin{aligned} &\left\| f_0 - \mathbb{E}[\hat{f}_h] \right\|_n^2 + \lambda \left[\left\| f_0 - \mathbb{E}[\hat{f}_h] \right\|_{H^1(\Omega)}^2 + \| g_0 - \mathbb{E}[\hat{g}_h] \|_{L^2(\Omega)}^2 \right] \\ &\leq C \Bigg\{ \inf_{(\varphi_h, p_h) \in \mathcal{W}_h} \left[\| f_0 - \varphi_h \|_n^2 + \lambda \| f_0 - \varphi_h \|_{H^1(\Omega)}^2 + \lambda \| g_0 - p_h \|_{L^2(\Omega)}^2 \right] \\ &+ \lambda \| g_0 \|_{L^2(\Omega)}^2 \Bigg\} \end{aligned} \tag{2.30}$$

for some constant $C > 0$ independent of h.

Proof. We set $f_h^* = \mathbb{E}[\hat{f}_h]$ and $g_h^* = \mathbb{E}[\hat{g}_h]$ and we recall that $\|\cdot\|_n$ is the norm induced by the bilinear form $l(\cdot,\cdot)$, i.e., $\|f\|_n^2 = l(f,f)$.
In order to prove Lemma 3 we can use the theory of saddle points systems. From equation (2.29) and the definition of $\mathcal{W}_h$ we have immediately

$$\begin{aligned} &\frac{1}{\lambda} l(\hat{f}_h^* - f_0, \psi_h) + a(\psi_h, \hat{g}_h^*) = 0 \;\; \forall \psi_h \in V^r_{h,\Gamma_D} \\ &a(\hat{f}_h^* - \varphi_h, v_h) = r(\hat{g}_h^* - p_h, v_h) \;\; \forall v_h \in V^r_{h,\Gamma_D}, \; (\varphi_h, p_h) \in \mathcal{W}_h. \end{aligned}$$

Choosing $(\varphi_h, p_h) \in \mathcal{W}_h$ we thus obtain

$$\begin{aligned} &\left\| \hat{f}_h^* - \varphi_h \right\|_n^2 + \lambda \| \hat{g}_h^* - p_h \|_{L^2(\Omega)}^2 = l(\hat{f}_h^* - \varphi_h, \hat{f}_h^* - \varphi_h) + \lambda r(\hat{g}_h^* - p_h, \hat{g}_h^* - p_h) \\ &= l(\hat{f}_h^* - f_0, \hat{f}_h^* - \varphi_h) + l(f_0 - \varphi_h, \hat{f}_h^* - \varphi_h) + \lambda r(\hat{g}_h^* - p_h, \hat{g}_h^* - p_h) \\ &= -\lambda a(\hat{f}_h^* - \varphi_h, \hat{g}_h^*) + l(f_0 - \varphi_h, \hat{f}_h^* - \varphi_h) + \lambda r(\hat{g}_h^* - p_h, \hat{g}_h^* - p_h) \\ &= l(f_0 - \varphi_h, \hat{f}_h^* - \varphi_h) - \lambda r(\hat{g}_h^* - p_h, p_h) \end{aligned} \tag{2.31}$$

since $\hat{f}_h^* - \varphi_h \in V^r_{h,\Gamma_D}$. We now bound the term $\hat{f}_h^* - \varphi_h$ in the H^1-norm using the coercivity of the bilinear form $a(\cdot,\cdot)$ in $H^1(\Omega)$:

$$\begin{aligned} \left\| \hat{f}_h^* - \varphi_h \right\|_{H^1(\Omega)}^2 &\leq \frac{1}{\alpha} a(\hat{f}_h^* - \varphi_h, \hat{f}_h^* - \varphi_h) = \frac{1}{\alpha} r(\hat{g}_h^* - p_h, \hat{f}_h^* - \varphi_h) \\ &\leq \frac{\sqrt{1 + C_P^2}}{\alpha} \| \hat{g}_h^* - p_h \|_{L^2(\Omega)} \left\| \hat{f}_h^* - \varphi_h \right\|_{H^1(\Omega)} \end{aligned}$$

where α is the coercivity constant and C_P is the constant in the Poincaré inequality (2.5), which holds thanks to Assumption 1. Summing this inequality to (2.31) we obtain

$$\begin{aligned}
&\left\|\hat{f}_h^* - \varphi_h\right\|_n^2 + \lambda\left[\frac{\alpha^2}{4(1+C_P^2)}\left\|\hat{f}_h^* - \varphi_h\right\|_{H^1(\Omega)}^2 + \|\hat{g}_h^* - p_h\|_{L^2(\Omega)}\right] \\
&\leq l(f_0 - \varphi_h, \hat{f}_h^* - \varphi_h) - \lambda r(\hat{g}_h^* - p_h, p_h) + \frac{\lambda}{4}\|\hat{g}_h^* - p_h\|_{L^2(\Omega)}^2 \\
&\leq \frac{1}{2}\|f_0 - \varphi_h\|_n^2 + \frac{1}{2}\left\|\hat{f}_h^* - \varphi_h\right\|_n^2 + \frac{\lambda}{2}\|\hat{g}_h^* - p_h\|_{L^2(\Omega)}^2 + 2\lambda\|g_0 - p_h\|_{L^2(\Omega)}^2 \\
&+ 2\lambda\|g_0\|_{L^2(\Omega)}^2 .
\end{aligned}$$

This inequality provides the bound

$$\begin{aligned}
&\left\|\hat{f}_h^* - \varphi_h\right\|_n^2 + \lambda\left[\left\|\hat{f}_h^* - \varphi_h\right\|_{H^1(\Omega)}^2 + \|\hat{g}_h^* - p_h\|_{L^2(\Omega)}^2\right] \\
&\leq C\left\{\|f_0 - \varphi_h\|_n^2 + \lambda\|g_0 - p_h\|_{L^2(\Omega)}^2 + \lambda\|g_0\|_{L^2(\Omega)}^2\right\}.
\end{aligned}$$

The final error bound (2.30) can now be obtained by triangular inequality and exploiting the arbitrariness of $(\varphi_h, p_h) \in \mathcal{W}_h$. □

We want now to split the error term on the constrained space $\mathcal{W}_h$ in two different errors for $\mathbb{E}[\hat{f}_h]$ and $\mathbb{E}[\hat{g}_h]$ on the space V_h^r. Assuming moreover that f_0 and g_0 are in proper Sobolev spaces $W^{s,p}(\Omega)$ we obtain the following result.

Theorem 3. *Using Finite Elements of degree r, if $f_0 \in W^{r+1,p}(\Omega)$ with $f_0|_{\Gamma_D} = h_D$ and $g_0 \in W^{r,p}(\Omega)$ with $g_0|_{\Gamma_D} = 0$ for $p > d$ then, under Assumption 3, there exists $h_0 > 0$ s.t. $\forall h \leq h_0$*

$$\begin{aligned}
&\left\|f_0 - \mathbb{E}[\hat{f}_h]\right\|_n^2 + \lambda\left[\left\|f_0 - \mathbb{E}[\hat{f}_h]\right\|_{H^1(\Omega)}^2 + \|g_0 - \mathbb{E}[\hat{g}_h]\|_{L^2(\Omega)}^2\right] \\
&\leq C\left[h^{2r}\left(|f_0|_{W^{r+1,p}(\Omega)}^2 + \lambda|g_0|_{W^{r,p}(\Omega)}^2\right) + \lambda\|g_0\|_{L^2(\Omega)}^2\right]. \qquad (2.32)
\end{aligned}$$

Proof. In order to prove the result we need to split in two parts the constrained error term

$$\inf_{(\varphi_h, p_h)\in\mathcal{W}_h}\left[\|f_0 - \varphi_h\|_n^2 + \lambda\|f_0 - \varphi_h\|_{H^1(\Omega)}^2 + \lambda\|g_0 - p_h\|_{L^2(\Omega)}^2\right]$$

in inequality (2.30).
We fix in the following $p_h \in V_{h,\Gamma_D}^r$ and we chose $\varphi_h \in V_h^r$ that satisfies $a(\varphi_h, v_h) = r(p_h, v_h) + F(v_h)$ and $\varphi_h|_{\Gamma_D} = h_{D,h}$, so that $(\varphi_h, p_h) \in \mathcal{W}_h$. Thanks to this choice we obtain the following bound

$$\|f_0 - \varphi_h\|_{H^1(\Omega)}^2 \leq C\left[\|f_0 - z_h\|_{H^1(\Omega)}^2 + \|g_0 - p_h\|_{L^2(\Omega)}^2\right] \qquad (2.33)$$

where z_h is an arbitrary function in V_h^r such that $z_h|_{\Gamma_D} = h_{D,h}$. This inequality is obtained thanks to the fact that

$$\|f_0 - \varphi_h\|_{H^1(\Omega)}^2 \leq 2\|f_0 - z_h\|_{H^1(\Omega)}^2 + 2\|z_h - \varphi_h\|_{H^1(\Omega)}^2$$

and that

$$\begin{aligned}\alpha \|z_h - \varphi_h\|^2_{H^1(\Omega)} &\le a(z_h - \varphi_h, z_h - \varphi_h) = a(f_0 - \varphi_h, z_h - \varphi_h) + a(z_h - f_0, z_h - \varphi_h) \\ &= r(g_0 - p_h, z_h - \varphi_h) + a(z_h - f_0, z_h - \varphi_h) \\ &\le \|g_0 - p_h\|_{L^2(\Omega)} \|z_h - \varphi_h\|_{L^2(\Omega)} + \|z_h - f_0\|_{H^1(\Omega)} \|z_h - \varphi_h\|_{H^1(\Omega)} \\ &\le C \left[\|g_0 - p_h\|_{L^2(\Omega)} + \|z_h - f_0\|_{H^1(\Omega)} \right] \|z_h - \varphi_h\|_{H^1(\Omega)} .\end{aligned}$$

The term $\|f_0 - \varphi_h\|^2_n$ can be bounded with the $W^{1,p}$-norm ($p > d$) of the same quantity, i.e.,

$$\|f_0 - \varphi_h\|_n \le C \|f_0 - \varphi_h\|_{W^{1,p}(\Omega)} . \tag{2.34}$$

We have in fact that

$$\frac{1}{n} \sum_{i=1}^{n} (f_0(\mathbf{p}_i) - \varphi_h(\mathbf{p}_i))^2 \le \max_{\mathbf{p}_i} (f_0(\mathbf{p}_i) - \varphi_h(\mathbf{p}_i))^2 \le \|f_0 - \varphi_h\|^2_{L^\infty}$$

and thanks to the Sobolev inequality (2.25) we obtain the upper bound (2.34).
We define now $f_{0h} \in V_h^r$ such that $a(f_{0h}, \psi_h) - r(g_0, \psi_h) = F(\psi_h)\ \forall \psi_h \in V^r_{h,\Gamma_D}$; the error term can be split in two parts

$$\|f_0 - \varphi_h\|_{W^{1,p}(\Omega)} \le \|f_0 - f_{0h}\|_{W^{1,p}(\Omega)} + \|f_{0h} - \varphi_h\|_{W^{1,p}(\Omega)} .$$

The first term on the right-hand side of the inequality represents the $W^{1,p}$-norm of the Finite Element error of the elliptic equation. The quantity f_{0h} can in fact be seen as the Finite Element approximation of the exact solution f_0 and for this reason (see, e.g., [2])

$$\|f_0 - f_{0h}\|_{W^{1,p}(\Omega)} \le C \inf_{\substack{z_h \in V_h^r \\ z_h|_{\Gamma_D} = h_{D,h}}} \|f_0 - z_h\|_{W^{1,p}(\Omega)} .$$

The second term of the right-hand side of the inequality can be bounded by

$$\|f_{0h} - \varphi_h\|_{W^{1,p}(\Omega)} \le C \|g_0 - p_h\|_{L^p(\Omega)}$$

where $p > d$. This bound is obtained thanks to the L^p-stability of the problem

$$a(f_{0,h} - \varphi_h, v_h) = r(g_0 - p_h, v_h)\ \ \forall v_h \in V^r_{h,\Gamma_D},$$

see, e.g., [2]. Therefore

$$\|f_0 - \varphi_h\|_n \le C \left(\inf_{\substack{z_h \in V_h^r \\ z_h|_{\Gamma_D} = h_{D,h}}} \|f_0 - z_h\|_{W^{1,p}(\Omega)} + \|g_0 - p_h\|_{L^p(\Omega)} \right) . \tag{2.35}$$

Collecting the bounds (2.33) and (2.35), since $L^p(\Omega) \subseteq L^2(\Omega)$ for $p \ge 2$ and Ω is bounded, we obtain for $p > d$, $d = 2, 3$ the unconstrained upper bound

$$\begin{aligned}&\left\| f_0 - \hat{f}_h^* \right\|^2_n + \lambda \left[\left\| f_0 - \hat{f}_h^* \right\|^2_{H^1(\Omega)} + \|g_0 - \hat{g}_h^*\|^2_{L^2(\Omega)} \right] \le \qquad (2.36)\\ &C \Bigg\{ \inf_{\substack{z_h \in V_h^r \\ z_h|_{\Gamma_D} = h_{D,h}}} \|f_0 - z_h\|^2_{W^{1,p}(\Omega)} + \lambda \inf_{p_h \in V^r_{h,\Gamma_D}} \|g_0 - p_h\|^2_{L^p(\Omega)} + \lambda \|g_0\|^2_{L^2(\Omega)} \Bigg\} .\end{aligned}$$

The classic error bound for the interpolant $\Pi_h^r v \in V_h^r$ of $v \in W^{r+1,p}(\Omega)$ with $p > 1$:

$$\|v - \Pi_h^r v\|_{W^{k,p}(\Omega)} \leq Ch^{r+1-k} |v|_{W^{r+1,p}(\Omega)} , \tag{2.37}$$

provides

$$\inf_{\substack{z_h \in V_h^r \\ z_h|_{\Gamma_D} = h_{D,h}}} \|f_0 - z_h\|_{W^{1,p}(\Omega)} \leq Ch^r |f_0|_{W^{r+1,p}(\Omega)}$$

$$\inf_{p_h \in V_{h,\Gamma_D}^r} \|g_0 - p_h\|_{L^p(\Omega)} \leq Ch^r |g_0|_{W^{r,p}(\Omega)} .$$

□

Notice that the inequality (2.32) can be split in two terms, the first term of the right-hand side goes to zero for $h \to 0$ while the second term $\|g_0\|_{L^2(\Omega)}^2$ is the same bias term obtained in the error splitting (2.22) and goes to zero when $\lambda \to 0$.

Remark 4. *In this thesis we propose an equal order Finite Element approximation for $\hat{f}$ and $\hat{g}$. Equal order Finite Elements are known to lead to sub-optimal convergence rates for the fourth order biharmonic problem (see, e.g., [4, 6]). However, here we are able to recover the optimal convergence rate thanks to the fact that the boundary conditions that are naturally associated to the smoothing problem are the same for $\hat{f}$ and $\hat{g}$. It should be noticed that the optimal convergence rate is recovered only if g_0 satisfies exactly the homogeneous Dirichlet boundary conditions on Γ_D, which might be a restrictive hypothesis. If g_0 does not satisfy the Dirichlet boundary conditions we should expect a "boundary term" decaying as $h^{1/2}$ both in two and three dimensions. Observe however that the approximation term $\lambda \inf_{p_h \in V_{h,\Gamma_D}^r} \|g_0 - p_h\|_{L^p(\Omega)}^2$ is always smaller than $\lambda \|g_0\|_{L^p(\Omega)}^2$ and for this reason the "boundary term" effect will be hidden by the bias term.*

2.7 Numerical simulations

2.7.1 Test 1

We propose to verify in a simple setting the convergence results shown in Section 2.6. We consider the bidimensional domain $\Omega = [0, 1] \times [0, 1]$ and we assume that the true underlying surface f_0 satisfies the following PDE

$$\begin{cases} \Delta f_0 = 2\,[x(x-1) + y(y-1)] & \text{in } \Omega \\ f_0 = 0 & \text{on } \partial\Omega \end{cases} \tag{2.38}$$

whose solution, $f_0 = xy(x - 1)(y - 1)$, is represented in Figure 2.1, Left. We consider the $n = 200$ observation points $\mathbf{p}_1, \ldots, \mathbf{p}_n$ represented in Figure 2.1, Right, and we want to test the convergence of $\|\mathbb{E}[\hat{f}_h - f_0]\|$ when $h \to 0$. For this reason we solve the estimation problem on different uniform structured meshes with size $h = 1/2, 1/4, \ldots, 1/2^9$.
We will consider different settings.

A. The observations are without noise, i.e., $z_i = f_0(\mathbf{p}_i)$, and the functional $J(f)$ penalizes the misfit of the governing PDE (2.38), i.e., $L = \Delta$ and $u = 2[x(x-1)+ y(y-1)]$.

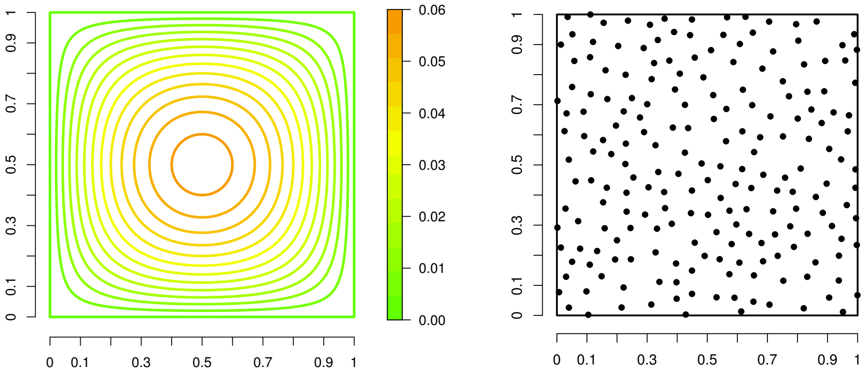

Figure 2.1: *Left: true surface* f_0 *used for the simulation studies of Test 1; the image displays the isolines* $(0, 0.005, 0.01, \ldots, 0.06)$. *Right: location points sampled uniformly on the domain for Test 1 and 2.*

B. The observations are without noise, i.e., $z_i = f_0(\mathbf{p}_i)$, but the functional $J(f)$ penalizes the misfit of a PDE different from the governing PDE (2.38). In particular $L = \Delta$ but $u \neq 2\,[x(x-1) + y(y-1)]$:

1. the penalized forcing term u is such that $g_0 = Lf_0 - u$ satisfies homogeneous Dirichlet boundary conditions on $\partial\Omega$: $u = 2(x(x-1)+y(y-1))(1+(x(x-1)y(y-1)))$;
2. different penalized forcing terms u are considered, such that $g_0 = Lf_0 - u$ does not satisfies homogeneous Dirichlet boundary conditions on $\partial\Omega$:
   (a) $u = (x(x-1) + y(y-1))$, which corresponds to the knowledge of the real forcing term up to a multiplying constant factor;
   (b) $u = 2x(x-1)$, which corresponds to the knowledge of only a part of the real forcing term;
   (c) $u = 2(x(x-1) + y(y-1)) + (x^{10} + y^{10} + (x-1)^{10} + (y-1)^{10})$, which forces g_0 to be equal to -1 on the boundary, with a relatively large boundary layer.

C. The observations are with noise, i.e., $z_i = f_0(\mathbf{p}_i) + \epsilon_i$, where $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$, and the functional $J(f)$ penalizes the misfit of the governing PDE (2.38), i.e., $L = \Delta$ and $u = 2\,[x(x-1) + y(y-1)]$.

Case A (no bias, no noise) We solve the estimation problem both with linear and quadratic Finite Elements fixing the roughness parameter $\lambda = 1$. We recall that we are using the same order of approximation for $\hat{f}_h$ and $\hat{g}_h$. The results of the linear and the quadratic mixed Finite Element approximation are shown respectively in the left and right panel of Figure 2.2. In particular we show the convergence of the error $|||\hat{f}_h - f_0|||$ as well as the convergence of each individual term of the norm $|||\cdot|||$, namely $\|\hat{f}_h - f_0\|_n$, $\|\hat{f}_h - f_0\|_{H^1(\Omega)}$ and $\|\hat{g}_h - g_0\|_{L^2(\Omega)}$. Since we are considering the case of observations without noise, $\mathbb{E}[\hat{f}_h - f_0] = \hat{f}_h - f_0$ and $\mathbb{E}[\hat{g}_h - g_0] = \hat{g}_h - g_0$. We notice that both

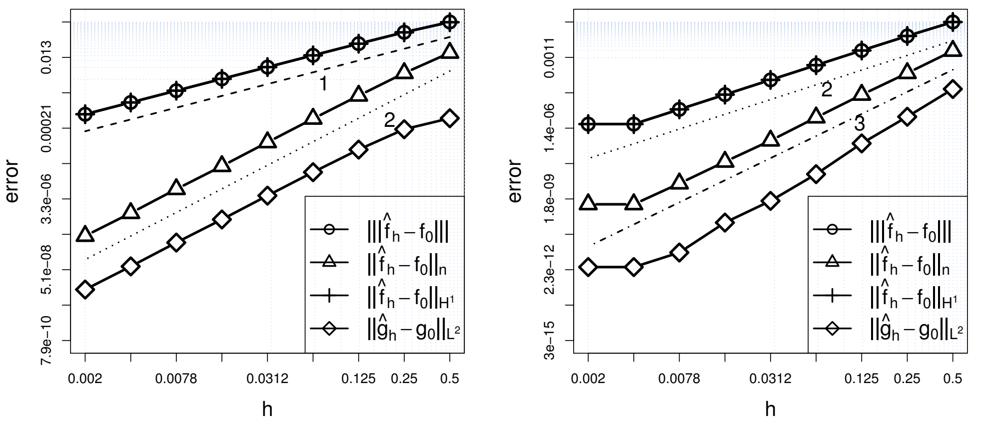

Figure 2.2: *Test 1, case A: convergence rates of the bias of the estimator in the norms $|||\hat{f}_h - f_0|||$, $\|\hat{f}_h - f_0\|_n$, $\|\hat{f}_h - f_0\|_{H^1(\Omega)}$ and $\|\hat{g}_h - g_0\|_{L^2(\Omega)}$ with $\lambda = 1$. Left: linear mixed Finite Element approximation. Right: quadratic mixed Finite Element approximation.*

with the linear and the quadratic approximation we obtain a rate of convergence equal to or higher than the expected rate for all the error terms. In particular, the H^1-norm of the error is the dominating term both in the linear and the quadratic approximation and decays as h in the case of linear Finite Elements and as h^2 in the case of quadratic Finite Elements. All the other terms are negligible. As expected, the norm $\|\cdot\|_n$ of $\hat{f}_h - f_0$ and the L^2-norm of $\hat{g}_h - g_0$ decay as h^2 in the case of linear Finite Elements and at least as h^3 in the case of quadratic Finite Elements.

Case B1 (bias with exact b.c., no noise) We solve the estimation problem with linear Finite Elements and we study the convergence for different values of the roughness parameter λ. Recall that, in this case, $g_0 = \Delta f_0 - u \neq 0$ satisfies the homogeneous Dirichlet boundary conditions. Figure 2.3 shows the rate of convergence of the error in different norms, when $\lambda = 0.05, 0.1, 0.2, 0.4$. As in case A, since the observations are without noise, $\mathbb{E}[\hat{f}_h - f_0] = \hat{f}_h - f_0$ and $\mathbb{E}[\hat{g}_h - g_0] = \hat{g}_h - g_0$. Notice that when the mesh is fine the approximation error in the norm $|||\cdot|||$ asymptotically approaches a value proportional to $\sqrt{\lambda}$, as expected from Theorem 3; this behavior is caused by the presence of the bias term in the error bound (2.32). The dominant term is in this case the L^2-norm of $\hat{g}_h - g_0$. This term has a different behavior for different values of λ: if λ is sufficiently small, it decays as h^2 before approaching the asymptote, otherwise it decays as h. It is thus necessary to use small values of λ in order to recover the expected convergence rate h^2 but even when using a large value of λ the rate of convergence of $|||\hat{f}_h - f_0|||$ is at least linear before reaching the saturation level caused by the bias term. The other two terms instead decay with the expected convergence rate for all the values of λ, before approaching the asymptote.

Case B2 (bias with wrong b.c., no noise) We consider three different forcing terms u such that $g_0 = \Delta f_0 - u \neq 0$ does not satisfy the homogeneous Dirichlet boundary conditions. In this case we study the error in the norm $|||\cdot|||$ over the whole domain Ω, which will

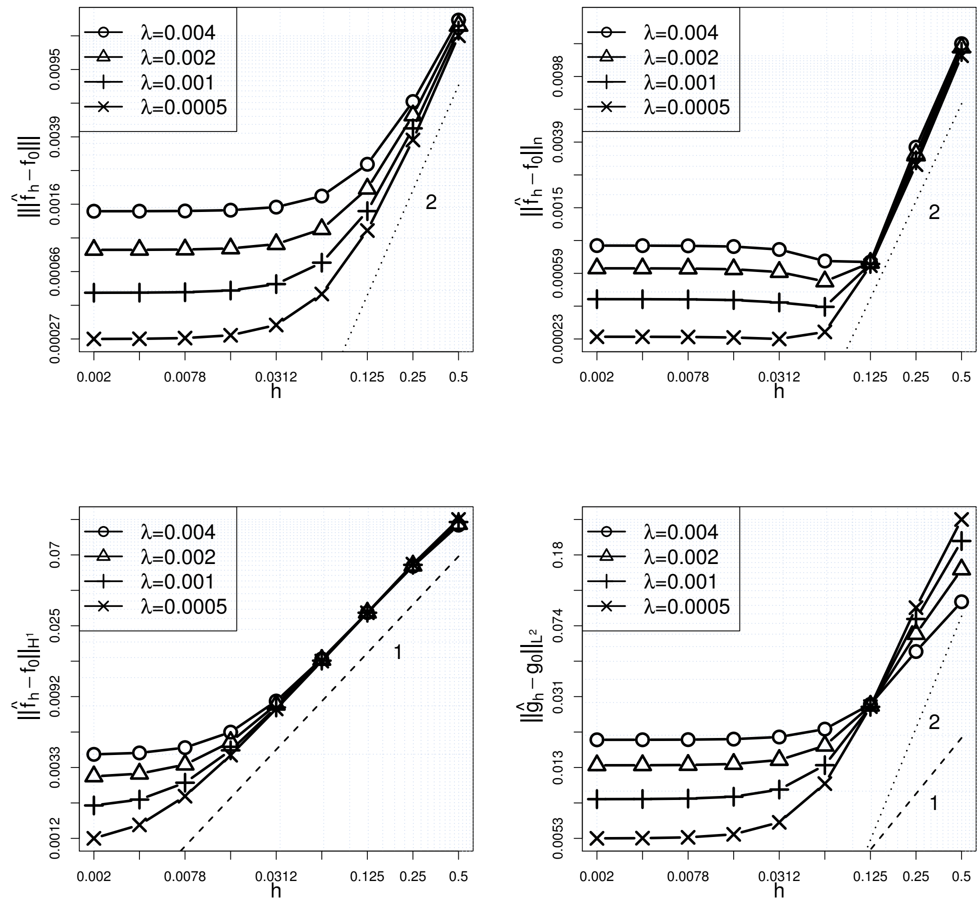

Figure 2.3: *Test 1, case B1: convergence rates of the bias of the estimator obtained with linear Finite Elements, when* $\lambda = 0.05, 0.1, 0.2, 0.4$. *Top left:* $\|\hat{f}_h - f_0\|$, *top right:* $\|\hat{f}_h - f_0\|_n$, *bottom left:* $\|\hat{f}_h - f_0\|_{H^1(\Omega)}$, *bottom right:* $\|\hat{g}_h - g_0\|_{L^2(\Omega)}$.

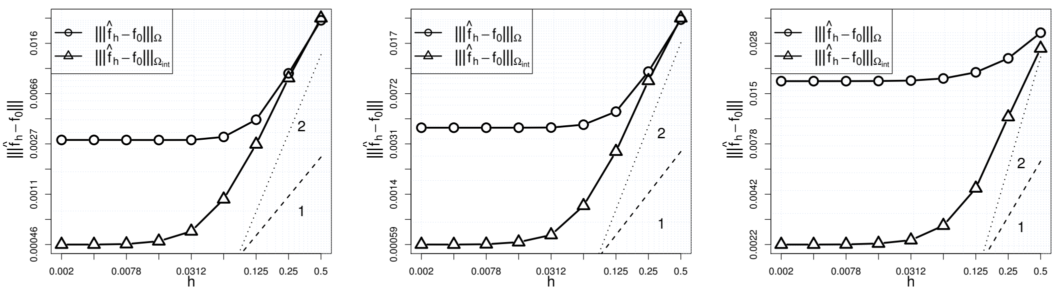

Figure 2.4: *Test 1, case B2: convergence rates of $\|\hat{f}_h - f_0\|_\Omega$ and $\|\hat{f}_h - f_0\|_{\Omega_{int}}$ using linear Finite Elements with $\lambda = 10^{-5}$. Left: case a), center: case b), right: c).*

be denoted by $\|\hat{f}_h - f_0\|_\Omega$, as well as over the subdomain $\Omega_{\text{int}} = [0.1, 0.9] \times [0.1, 0.9]$, denoted by $\|\hat{f}_h - f_0\|_{\Omega_{\text{int}}}$. As highlighted in Remark 4, the former error should be affected by a "boundary term" decaying as $h^{1/2}$, while the latter does not include the error at the boundary. As in case A and B1, since the observations are without noise, $\mathbb{E}[\hat{f}_h - f_0] = \hat{f}_h - f_0$ and $\mathbb{E}[\hat{g}_h - g_0] = \hat{g}_h - g_0$. The results obtained with the three forcing terms are represented respectively in the left, center and right panels of Figure 2.4. In theory we would expect a different rate of convergence for the two errors, which should be more clearly visible when the mesh is fine. On the other hand the numerical simulations do not display any significant difference between the convergence rates of the two errors in all the three cases; this is due to the presence of the bias, which is asymptotically approached by both the error terms, that hides the expected convergence rate. Thus, using a forcing term such that g_0 does not satisfy the homogeneous Dirichlet boundary conditions, does not affect too much the surface estimation.

Case C (no bias, with noise) We add some noise to the pointwise evaluations $f_0(\mathbf{p}_i)$ of the surface: for each location point we sample independent errors, $\epsilon_1, \ldots, \epsilon_n$, from a zero mean Gaussian distribution $\mathcal{N}(0, \sigma^2)$, with different standard deviations $\sigma = 0.005, 0.01, 0.02$. The first value of σ corresponds to a rather high signal to noise ratio, since the value of the true surface varies from 0 to 0.062, while the last corresponds to a very low signal to noise ratio with errors of the same order of magnitude as the variation of f_0. The values z_i, obtained from model (2.1), are shown in Figure 2.5. We can notice that the observations with small additive noise, represented in the left panel of Figure 2.5, are similar to the evaluation of f_0 in the sampling points, while the observations with large errors, represented in the right panel of the same figure, are far from the true underlying surface. The results obtained solving the estimation problem with linear Finite Elements, fixed roughness parameter $\lambda = 0.01$ and the exact PDE penalized are shown in Figure 2.6. The represented results concern a single replicate of the experiment and show a typical behavior of the error convergence in different norms. Notice that in Figure 2.6 the represented errors include both the approximation error and the error associated to the noisy observations. Due to the presence of noise, both $\|\hat{f}_h - f_0\|$ and $\|\hat{f}_h - f_0\|_n$ reach quite soon a saturation limit proportional to the standard deviation of the noise σ. Refining further the mesh still provides better approximation of the first derivatives as shown by the convergence of $\|\hat{f}_h - f_0\|_{H^1(\Omega)}$.

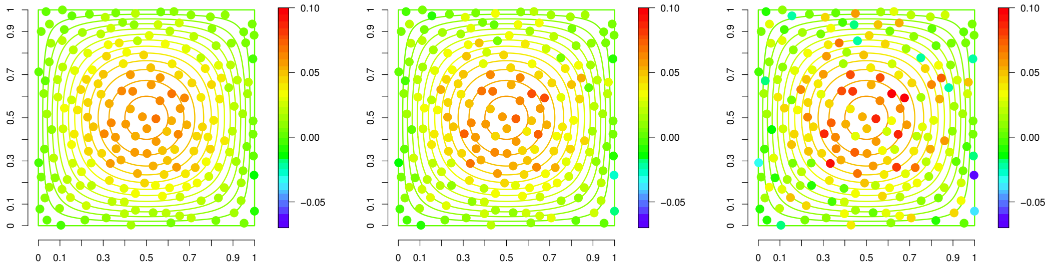

Figure 2.5: *Test 1, case C: value of the observations $z_1, \ldots, z_n$ obtained from model* (2.1) *adding noise with different values of standard deviation σ, superimposed to the true underlying surface f_0 (the image displays the isolines $(0, 0.005, 0.01, \ldots, 0.06)$). Left: $\sigma = 0.005$, center: $\sigma = 0.01$, right: $\sigma = 0.02$.*

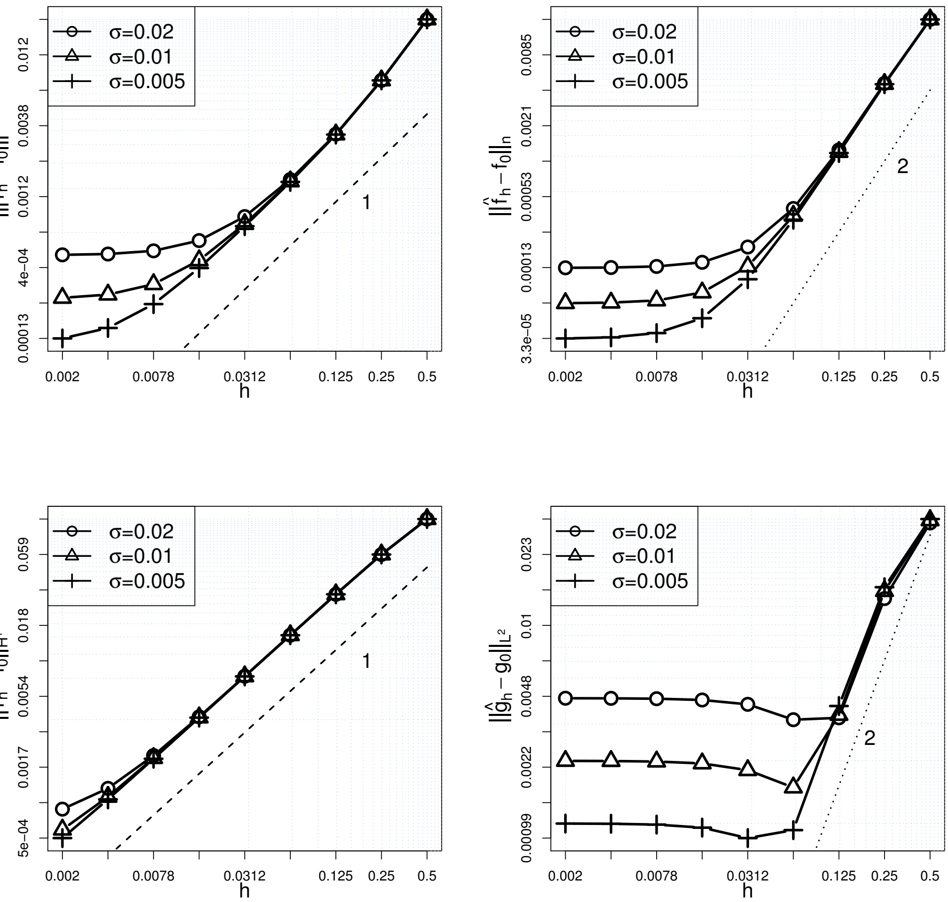

Figure 2.6: *Test 1, case C: convergence rates of the bias of the estimator obtained with linear Finite Elements and $\lambda = 1$, when the error in the observations is generated form a Gaussian distribution with different standard deviations $\sigma = 0.005, 0.01, 0.02$. Top left: $\|\hat{f}_h - f_0\|$, top right: $\|\hat{f}_h - f_0\|_n$, bottom left: $\|\hat{f}_h - f_0\|_{H^1(\Omega)}$, bottom right: $\|\hat{g}_h - g_0\|_{L^2(\Omega)}$.*

2.7.2 Test 2

We test the convergence also in a different simulation study concerning a diffusion-transport-reaction (DTR) PDE. We consider the domain $\Omega = [0,1] \times [0,1]$ and we assume that the true underlying field f_0 satisfies the following PDE

$$\begin{cases} Lf_0 = -1 & \text{in } \Omega \\ f_0 = 0 & \text{on } \partial\Omega \end{cases} \tag{2.39}$$

where the operator L is the diffusion-transport-reaction operator defined in (2.3) with parameters $\mathbf{K}_{11} = 4$, $\mathbf{K}_{22} = 1$, $\mathbf{K}_{12} = \mathbf{K}_{21} = 0$, $\mathbf{b}_1 = 2$, $\mathbf{b}_2 = 1$ and $c = 1$, $\mathbf{K}_{ij}$ and $\mathbf{b}_i$ being respectively the element (i,j) of the diffusion tensor $\mathbf{K}$ and the i-th element

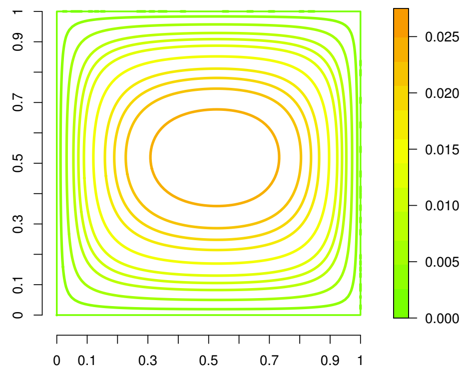

Figure 2.7: *True surface f_0 used for the simulation study of Test 2; the image displays the isolines $(0, 0.0025, 0.005, \ldots, 0.03)$.*

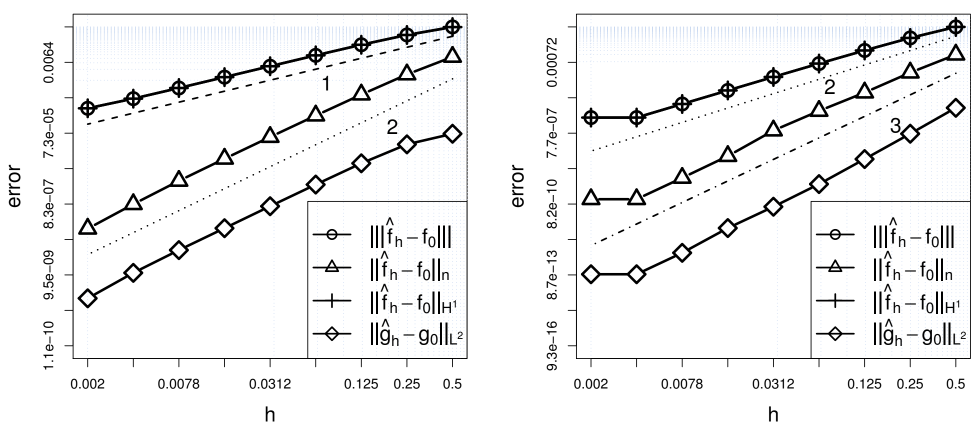

Figure 2.8: *Test 2: convergence rates of the bias of the estimator in the norms $|||\hat{f}_h - f_0|||$, $\|\hat{f}_h - f_0\|_n$, $\|\hat{f}_h - f_0\|_{H^1(\Omega)}$ and $\|\hat{g}_h - g_0\|_{L^2(\Omega)}$ with $\lambda = 1$. Left: linear mixed Finite Element approximation. Right: quadratic mixed Finite Element approximation.*

of the transport vector $\mathbf{b}$. The solution of the PDE (2.39) is represented in Figure 2.7. We consider the $n = 200$ observation points $\mathbf{p}_1, \ldots, \mathbf{p}_n$ represented in Figure 2.1, Right. In this case we only show the convergence of $\|\mathbb{E}[\hat{f}_h - f_0]\| = \|\hat{f}_h - f_0\|$ when the observations are without noise and the functional $J(f)$ penalizes the misfit of the governing PDE (2.39). The results obtained solving the estimation problem with linear and quadratic Finite Elements and $\lambda = 1$ are shown in Figure 2.8. We can notice that also in the DTR case we obtain a rate of convergence equal to or higher than the expected rate for all the error terms both with the linear and the quadratic approximation. The H^1-norm is still the dominating term while all the other terms are negligible. As in the Laplacian case, the error terms $\|\hat{f}_h - f_0\|_n$ and $\|\hat{g}_h - g_0\|_{L^2(\Omega)}$ decay as h^2 for linear Finite Elements and at least as h^3 for quadratic Finite Elements.

2.7.3 Test 3

In the last simulation study we test the error convergence in a setting similar to the velocity field estimation problem and the simulation studies presented in Chapter 1. We consider in particular a circular domain centered in zero, with unitary radius, $\Omega = B_1$, which is similar to the almost circular artery cross-section used in the blood velocity field estimation problem presented in Chapter 1. We assume that the true underlying field f_0 satisfies the following PDE

$$\begin{cases} \Delta f_0 = y & \text{in } \Omega \\ f_0 = 0 & \text{on } \partial\Omega \end{cases} \tag{2.40}$$

whose solution $f_0 = y/8(1-x^2+y^2)$ is represented in Figure 2.9, Left. We consider the $n = 200$ observation points $\mathbf{p}_1, \ldots, \mathbf{p}_n$, represented in Figure 2.9, Right, on the circular domain. As in the previous simulation study, in this case we only show the convergence when the observations are without noise and the operator penalized in the functional $J(f)$ is the governing PDE (2.40). The results obtained solving the estimation problem

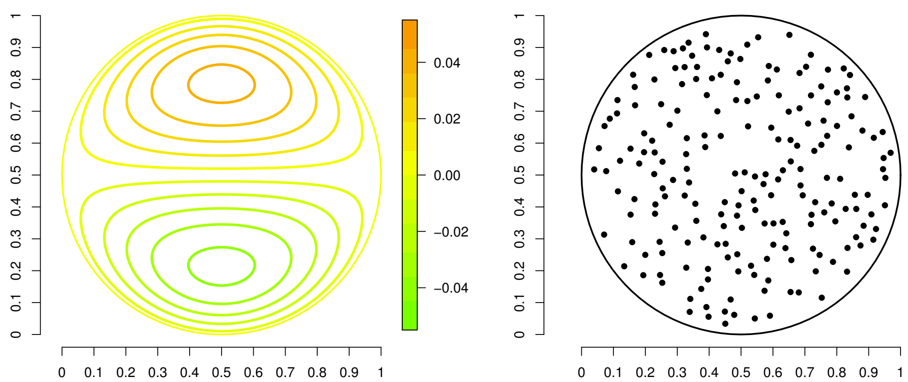

Figure 2.9: *Left: true surface f_0 used for the simulation study of Test 3; the image displays the isolines* $(-0.055, -0.045, -0.035, \ldots, 0.055)$. *Right: location points sampled uniformly on the domain for Test 3.*

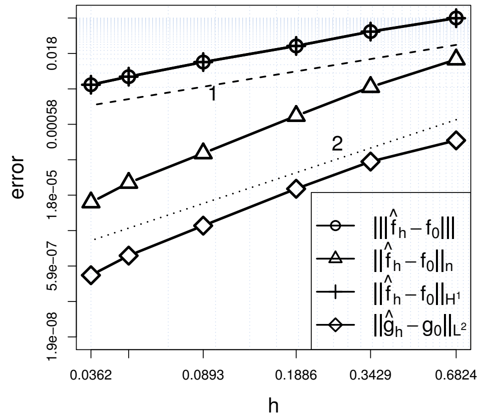

Figure 2.10: *Test 3: convergence rates of the bias of the estimator in the norms* $\|\hat{f}_h - f_0\|$, $\|\hat{f}_h - f_0\|_n$, $\|\hat{f}_h - f_0\|_{H^1(\Omega)}$ *and* $\|\hat{g}_h - g_0\|_{L^2(\Omega)}$ *with* $\lambda = 1$ *and linear mixed Finite Element approximation.*

with linear Finite Elements and $\lambda = 1$ on 6 different uniform unstructured meshes are shown in Figure 2.10. As expected, the rate of convergence of the error $|||\mathbb{E}[\hat{f}_h - f_0]||| = |||\hat{f}_h - f_0|||$ is h, since $\|\hat{f}_h - f_0\|_{H^1(\Omega)}$ is the dominating term, while $\|\hat{f}_h - f_0\|_n$ and $\|\hat{g}_h - g_0\|_{L^2(\Omega)}$ decay as h^2. We thus obtain an optimal convergence rate of the error also in the case of a circular domain.

2.8 Surface estimator for areal data

The field smoothing method presented in Section 2.2 can be extended to the case of areal data that represent quantities computed on some subregions. This is useful in many applications of interest and it is for instance the case of the estimation of blood velocity field from Echo-Doppler data presented in Chapter 1; the Echo-Doppler data represent in fact the mean velocity of the blood cells on a subdomain within an artery and cannot be approximated with pointwise observations.
Let $D_i \subset \Omega$, for $i = 1, \ldots, N$, be some subdomains and $\bar{z}_i$, for $i = 1, \ldots, N$, the mean value of a quantity of interest on the subdomains. We consider the following model for the observations $\bar{z}_i$:

$$\bar{z}_i = \frac{1}{|D_i|} \int_{D_i} f_0 + \eta_i. \tag{2.41}$$

The error terms η_i have zero mean and variances $\bar{\sigma}_i^2$. The variances $\bar{\sigma}_i^2$ depend inversely on the dimension of the beams D_i under the assumption that the number of observations in a subdomain is proportional to its dimension. As, we have seen in Chapter 1, this model can be derived starting from the model for pointwise observations.
Similarly to Chapter 1, in order to estimate the field we minimize the penalized sum-of-square-error functional

$$\bar{J}(f) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{|D_i|} \left(\int_{D_i} (f - \bar{z}_i)\, d\mathbf{p} \right)^2 + \lambda \int_{\Omega} (Lf - u)^2 \tag{2.42}$$

over the space V, defined in Section 2.2. The first term is a weighted least-square-error functional for areal data over subdomains D_i, weighted with the inverse of the variances $\bar{\sigma}_i^2$, under the assumption that $\bar{\sigma}_i^2 \propto 1/|D_i|$.
As in the pointwise data framework, the functional $\bar{J}$ is slightly different from the functional used in Chapter 1 for areal data. Specifically, the least square term is divided by N, which corresponds to multiplying N to the roughness parameter λ in the functional (1.8).

Existence and uniqueness of the estimator $\hat{f} = \text{argmin}_{f\in V} \bar{J}(f)$ is provided by the following theorem.

Theorem 4. *The estimator $\hat{f}$ exists, is unique and is obtained solving the system of PDEs:*

$$\begin{cases} L\hat{f} = u + \hat{g} & \text{in } \Omega \\ \mathcal{B}_c\hat{f} = h & \text{on } \partial\Omega \end{cases} \qquad \begin{cases} L^*\hat{g} = -\frac{1}{N\lambda}\sum_{i=1}^N \frac{1}{|D_i|}\mathbb{I}_{D_i}\int_{D_i}(\hat{f} - \bar{z}_i) & \text{in } \Omega \\ \mathcal{B}_c{}^*\hat{g} = 0 & \text{on } \partial\Omega \end{cases} \tag{2.43}$$

where $\hat{g} \in \mathcal{G}$ represents the misfit of the penalized PDE, L^* is the adjoint operator of L, described by equation (2.13)*, and $\mathcal{B}_c{}^*$ is the operator that defines the boundary conditions of the adjoint problem, summarized in* (2.14).

The proof is analogous to the proof of Theorem 2. The existence and uniqueness of the estimator is in fact obtained, thanks to Theorem 1, writing the functional $\bar{J}(f)$ as the quadratic form (2.8). The proof of the well posedness of the problem in the areal case is easier than the one presented in Section 2.3 and it is similar to classical results in control theory. Data are in fact distributed and it's not necessary to require more regularity as in the case of punctual observations.

The estimator is then discretized by means of the mixed Finite Element method described in Section 2.5. The Finite Element estimator $\hat{f}_h$ can be written as $\hat{f}_h = \hat{\mathbf{f}}^T\boldsymbol{\psi} + \hat{\mathbf{f}}_D^T\boldsymbol{\psi}^D$ where $\hat{\mathbf{f}}$ is the solution of the linear system

$$\begin{bmatrix} \bar{\mathbf{\Psi}}^T\mathbf{W}\bar{\mathbf{\Psi}}/(N\lambda) & \mathbf{A}^T \\ \mathbf{A} & -\mathbf{R} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{\Psi}}^T\mathbf{W}\bar{\mathbf{z}}/(N\lambda) - \bar{\mathbf{\Psi}}^T\bar{\mathbf{\Psi}}^D\mathbf{f}_D/(N\lambda) \\ \mathbf{u} + \mathbf{h}_N + \mathbf{h}_R - \mathbf{A}^D\mathbf{f}_D \end{bmatrix}. \tag{2.44}$$

where $\bar{\mathbf{\Psi}}_{ik} = 1/|D_i|\int_{D_i}\psi_k$ and $\bar{\mathbf{\Psi}}_{ik}^D = 1/|D_i|\int_{D_i}\psi_k^D$ represents the spatial average of the basis functions on the subdomains D_i, $\mathbf{W} = \text{diag}(|D_1|, \dots, |D_N|)$ is the weight matrix and $\bar{\mathbf{z}} = (\bar{z}_1, \dots, \bar{z}_N)^T$ is the vector of mean values on subdomains.

As in the case of pointwise observations we can obtain a bound for the bias of the estimator $\hat{f}$ that corresponds exactly to the bound (2.22). We can use the results obtained in the pointwise case also for the study of the convergence of the bias of the Finite Element estimator to the true underlying surface f_0. In the areal data case we can also relax the hypothesis on f_0 and g_0 in Theorem 3.

Theorem 5. *Using Finite Elements of degree r, if $f_0 \in H^{r+1}(\Omega)$ with $f_0|_{\Gamma_D} = h_D$ and $g_0 \in H^r(\Omega)$ with $g_0|_{\Gamma_D} = 0$, we obtain, under Assumption 1,*

$$\frac{1}{N}\sum_{i=1}^N \frac{1}{|D_i|}\int_{D_i}(f_0 - \mathbb{E}[\hat{f}^*])^2 + \lambda\left[\left\|f_0 - \mathbb{E}[\hat{f}_h]\right\|_{H^1(\Omega)}^2 + \|g_0 - \mathbb{E}[\hat{g}_h]\|_{L^2(\Omega)}^2\right]$$
$$\leq C\left[h^{2r}\left(|f_0|_{H^{r+1}(\Omega)}^2 + |g_0|_{H^r(\Omega)}^2\right) + \lambda\|g_0\|_{L^2(\Omega)}^2\right]. \tag{2.45}$$

Proof. If we define $\hat{f}_h^* = \mathbb{E}[\hat{f}_h]$, $\hat{g}_h^* = \mathbb{E}[\hat{g}_h]$ and the bilinear form $l(\cdot,\cdot)$ as

$$l(f,\psi) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{|D_i|}\int_{D_i} f\psi$$

we can easily obtain the bound (2.30) for the norm of the bias defined as:

$$\left|\left|\left|\hat{f}_h^* - f_0\right|\right|\right| = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{|D_i|}\int_{D_i}(\hat{f}_h^* - f_0)^2 + \lambda\left[\left\|\hat{f}_h^* - f_0\right\|^2_{H^1(\Omega)} + \|\hat{g}_h^* - g_0\|^2_{L^2(\Omega)}\right].$$

Since the norm associated to the bilinear form $l(\cdot,\cdot)$ is bounded by the H^1-norm we have that

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{|D_i|}\int_{D_i}(f_0 - \hat{f}_h^*)^2 + \lambda\left[\left\|f_0 - \hat{f}_h^*\right\|^2_{H^1(\Omega)} + \|g_0 - \hat{g}_h^*\|^2_{L^2(\Omega)}\right]$$
$$\leq C\left\{\inf_{(\varphi_h,p_h)\in\mathcal{W}_h}\left[\|f_0-\varphi_h\|^2_{H^1(\Omega)} + \lambda\|g_0 - p_h\|^2_{L^2(\Omega)}\right] + \lambda\|g_0\|^2_{L^2(\Omega)}\right\}.$$

The inequality (2.33) still holds for $(\varphi_h, p_h)\in\mathcal{W}_h$ and $z_h \in V_h^r$ and we obtain

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{|D_i|}\int_{D_i}(f_0 - \hat{f}_h^*)^2 + \lambda\left[\left\|f_0 - \hat{f}_h^*\right\|^2_{H^1(\Omega)} + \|g_0 - \hat{g}_h^*\|^2_{L^2(\Omega)}\right] \leq$$
$$C\left\{\inf_{\substack{z_h\in V_h^r\\ z_h|_{\Gamma_D}=h_{D,h}}}\|f_0 - z_h\|^2_{H^1(\Omega)} + \lambda\inf_{p_h\in V^r_{h,\Gamma_D}}\|g_0 - p_h\|^2_{L^2(\Omega)} + \lambda\|g_0\|^2_{L^2(\Omega)}\right\}$$

Using the classic error bound (2.37) we obtain the desired result. □

CHAPTER 3

A first approach to time-dependent Spatial Regression with PDE penalization

3.1 Introduction

In this chapter we extend Spatial Regression with PDE penalization (SR-PDE) in order to model surfaces evolving in time. Such extension is particularly interesting for the velocity field estimation problem presented in Chapter 1. In fact, starting from spatially located echo doppler signals, this extension allows to study how the blood-velocity field varies during the time of the heartbeat. As explained in Chapter 1, the echo doppler scan provides a time-dependent signal that represents an histogram of the measured velocities, evolving in time. Figure 3.1 shows part of the ECD signal registered in the

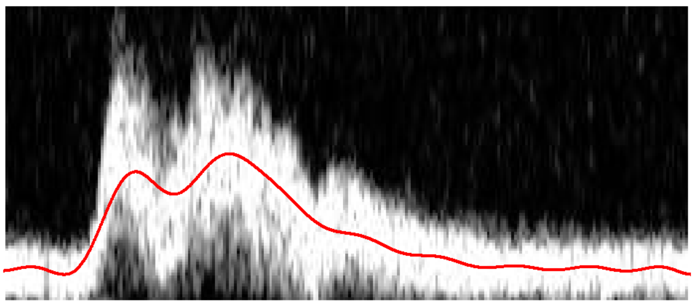

Figure 3.1: *Part of the echo doppler signal with superimposed the estimated time-dependent mean velocity registered within the corresponding beam.*

central beam of the carotid cross-section located 2 cm before the bifurcation and, in red, the corresponding time-dependent mean velocity on the beam, estimated from the entire ECD signal by means of a Fourier smoothing.

Following the approach presented in the previous chapters, we propose to estimate time evolving surfaces minimizing a penalized sum-of-square-error functional, with the penalty term involving a time-dependent partial differential operator. A similar approach for space-time smoothing has been recently developed in [1] and [24]. The method proposed in these two works generalizes soap-film smoothing by means of a tensor product smoothing of soap-film smoothing basis in space and univariate splines in time. This approach corresponds to the penalization of the two marginal roughness terms, which involve the Laplace operator in space and the second derivative in time.

Instead of using a tensor product approach, we propose to penalize the misfit of a time-dependent PDE modeling the phenomenon under study, $\dot{f}_0 + Lf_0 = u$, where $\dot{f}_0$ is the time derivative of f_0 and L is a differential operator in space. We consider square integrable observations in time and both the case of pointwise and areal data in space. The estimation problems are well posed and can be discretized in space by means of the Finite Element method, similarly to SR-PDE described in Chapter 1, and in time by means of the Finite Difference method.

The chapter is organized as follows. Section 3.2 extends SR-PDE to the case of dynamic surfaces when the observations are pointwise in space and functional in time. Section 3.3 proves the well-posedness of the estimation problem. Section 3.4 describes the discretization of the estimation problem by means of the Finite Element method in space and the Finite Difference method in time. Section 3.5 extends the models to the case of areal data in space, particularly interesting for the analysis of ECD measurements. Section 3.6 shows a simple simulation study for the estimation of a dynamic surface starting from pointwise data in space and functional data in time. Section 3.7 presents the results obtained with time-dependent SR-PDE in the blood velocity field estimation within the MACAREN@MOX project.

3.2 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 time-dependent observations $z_i(t)$, for $i = 1, \ldots, n$, located at points $\mathbf{p}_i = (x_i, y_i) \in \Omega$ with $t \in [0, T]$. We assume that the observations $z_i(t)$ are square integrable random processes with mean

$$\mathbb{E}\left[z_i(t)\right] = f_0(\mathbf{p}_i, t) \tag{3.1}$$

where $f_0 : \Omega \times [0, T] \to \mathbb{R}$ is the surface that we want to estimate. In our application Ω will be the carotid cross-section, $z_i(t)$ will represent the velocity measured in the longitudinal direction of the artery during the heartbeat of period T and the surface f_0 will represent the velocity field varying in time.

We assume to have a prior knowledge on the phenomenon under study that can be described in terms of a time-dependent Partial Differential Equation (PDE), $\dot{f}_0 + Lf_0 = u$, where $\dot{f}_0$ is the time derivative of f_0. Moreover, we assume that f_0 has to satisfy some boundary conditions on $\partial\Omega$ and an initial condition at time $t = 0$. As explained in the previous chapters we want to take advantage of the prior knowledge on the phenomenon in the surface estimation. To this end, we generalize SR-PDE presented in Chapter 1 so that the roughness penalty may include the misfit of the time-dependent partial differential operator known to model the phenomenon under study. In particular, we propose to estimate the surface f_0 by minimizing the penalized sum-of-square-error functional

$$J_T(f) = \sum_{i=1}^{n} \int_0^T (f(\mathbf{p}_i, t) - z_i(t))^2 + \lambda \int_0^T \int_\Omega \left(\frac{\partial f}{\partial t} + Lf - u \right)^2 \tag{3.2}$$

with respect to $f \in V_T$, where V_T is the space of functions $f \in L^2(0, T; H^2(\Omega))$ with first derivative in time $\dot{f} \in L^2(0, T; L^2(\Omega))$ that moreover satisfy the imposed boundary conditions on $\partial\Omega$ and the initial condition at time $t = 0$. The space $L^2(0, T; H)$ is defined as the space of functions such that

$$\int_0^T \|f(t)\|_H^2 \, dt < +\infty.$$

The space V_T contains thus the functions such that

$$\int_0^T \|f(t)\|_{H^2(\Omega)}^2 \, dt + \int_0^T \left\| \dot{f}(t) \right\|_{L^2(\Omega)}^2 dt < +\infty.$$

As in SR-PDE models described in Chapter 1, the functional J_T is composed by a data fitting criterion, consisting in the sum of square errors now integrated in time, and a model fitting criterion, formalized as a penalizing term. The PDE penalized in the roughness term is a parabolic PDE supposed to model the phenomenon under study:

$$\begin{cases} \dfrac{\partial f_0}{\partial t} + Lf_0 = \tilde{u} & \text{in } \Omega \times (0, T) \\ f_0(\mathbf{x}, 0) = s(\mathbf{x}) & \text{in } \Omega \\ \mathcal{B}_c f_0 = h & \text{on } \partial\Omega \times (0, T] \end{cases} \tag{3.3}$$

where the operator L, defined in equation (1.3), is the same elliptic operator that is penalized in Chapter 1. The parameters of the PDE can be space-varying on Ω but not time varying; i.e., $\mathbf{K} = \mathbf{K}(\mathbf{x})$, $\mathbf{b} = \mathbf{b}(\mathbf{x})$ and $c = c(\mathbf{x})$. We do not consider the case of time-varying parameters since they would reduce the regularity of the solution.

The forcing term $\tilde{u} = u + g_0 \in L^2(0, T; L^2(\Omega))$ is composed by a known and fixed part u and an unknown term, called g_0, that will be estimated from data. The initial value $s \in H^1(\Omega)$ and $\tilde{u}$ have to satisfy some joint conditions in order not to reduce the regularity of the solution.

The boundary conditions of the PDE are homogeneous or non-homogeneous Dirichlet, Neumann, Robin (or mixed) conditions, and they are summarized in

$$\mathcal{B}_c f_0 = \begin{cases} f_0 & \text{on } \Gamma_D \times (0, T] \\ \mathbf{K}\nabla f_0 \cdot \boldsymbol{\nu} & \text{on } \Gamma_N \times (0, T] \\ \mathbf{K}\nabla f_0 \cdot \boldsymbol{\nu} + \gamma f_0 & \text{on } \Gamma_R \times (0, T] \end{cases} \qquad h = \begin{cases} h_D & \text{on } \Gamma_D \times (0, T] \\ h_N & \text{on } \Gamma_N \times (0, T] \\ h_R & \text{on } \Gamma_R \times (0, T] \end{cases} \tag{3.4}$$

where $\boldsymbol{\nu}$ is the outward unit normal vector to $\partial\Omega$, $\gamma \in \mathbb{R}$ is a positive constant and $\partial\Omega = \bar{\Gamma}_D \cup \bar{\Gamma}_N \cup \bar{\Gamma}_R$, with $\Gamma_D, \Gamma_N, \Gamma_R$ not overlapping. In the following we suppose also that Assumption 1 holds.

The functional $J_T(f)$ is well defined if $f \in V_T$ since this space contains functions continuous in space and square integrable in time, such that

$$\int_0^T \left(\sup_{\mathbf{x} \in \bar{\Omega}} f(\mathbf{x}, t) \right)^2 dt < \infty$$

thanks to the embedding $H^2(\Omega) \subset C(\bar{\Omega})$ if $\Omega \subset \mathbb{R}^d$ with $d \leq 3$. Both the penalty term and the least square term are thus well defined for functions in V_T. For data in $\mathbb{R}^d$ with $d > 3$ one has to require more regularity in order to obtain $f \in L^2(0, T; C(\bar{\Omega}))$; in particular one needs $f \in L^2(0, T; H^s(\Omega))$ with $s > d/2$; see, e.g., [12].

The estimation problem is formulated as follows.

Problem 5. *Find $\hat{f} \in V_T$ such that*

$$\hat{f} = \operatorname*{argmin}_{f \in V_T} J_T(f).$$

The estimation problem is well posed if we assume some regularity on the parameters of the PDE and on the domain Ω. In particular, in the case $d \leq 3$, we make the following assumption.

Assumption 4. *The parameters of the PDE are such that $\forall \tilde{u} \in L^2(0, T; L^2(\Omega))$ there exists a unique solution f_0 of the PDE (3.3), which moreover satisfies $f_0 \in L^2(0, T; H^2(\Omega)) \cap L^\infty(0, T; H^1_{\Gamma_D}(\Omega))$ and $\dot{f}_0 \in L^2(0, T; L^2(\Omega))$.*

3.3 Estimation problem

To analyze the existence and the uniqueness of the solution of Problem 5, we introduce a new quantity $g \in \mathcal{G}_T = L^2(0, T; L^2(\Omega))$, which is defined as $g = \dot{f} + Lf - u$, where L is the second order elliptic operator (1.3)
We also introduce the space $V_{T,0} = \{v \in V_T : \mathcal{B}_c v = 0 \text{ and } v(\mathbf{x}, 0) = 0\}$, which represents the space of functions in V_T with homogeneous boundary conditions and homogeneous initial value, and the operator $B_T : L^2(0, T; L^2(\Omega)) \to V_{T,0}$, such that $B_T \tilde{u}$ is the unique solution of the PDE (3.3) with forcing term $\tilde{u}$ and homogeneous boundary conditions, i.e., $L(B_T \tilde{u}) = \tilde{u}$ in $\Omega \times (0, T)$, $B_T \tilde{u}(\mathbf{x}, 0) = 0$ in Ω and $\mathcal{B}_c(B_T \tilde{u}) = 0$ on $\partial\Omega \times (0, T]$. Under Assumptions 1 and 4, thanks to the well-posedness and the regularity of the PDE (3.3), the operator B_T is an isomorphism between $L^2(0, T; L^2(\Omega))$ and $V_{T,0}$, and the following inequality holds

$$\int_0^T \|B_T \tilde{u}(t)\|^2_{H^2(\Omega)} \leq C \int_0^T \|\tilde{u}(t)\|^2_{L^2(\Omega)} . \tag{3.5}$$

The solution of the PDE (3.3) can thus be written as $f = f_b + B_T \tilde{u}$ where f_b is the solution of the PDE with homogeneous forcing term, non-homogeneous initial value and non-homogeneous boundary conditions.

The existence and uniqueness of the estimator is stated in the following theorem.

Theorem 6. *Under Assumptions 1 and 4, the solution of Problem 5 exists and is unique.*

Proof. Thanks to the definition of g we can write f as an affine transformation of g, i.e., $f = f_b + B_T(u+g)$, and the functional (3.2) as

$$J_{T,g}(g) = J_T(f_b + B_T(u+g)) = \sum_{i=1}^{n} \int_0^T (B_T(u+g)(\mathbf{p}_i,t) + f_b(\mathbf{p}_i,t) - z_i(t))^2 + \lambda \int_0^T \|g(t)\|^2_{L^2(\Omega)}. \tag{3.6}$$

This reformulation of the functional J_T is very useful since we can now write $J_{T,g}$ in the quadratic form (2.8) of Theorem 1 where

$$\begin{aligned} \mathcal{A}(g,\varphi) &= \sum_{i=1}^{n} \int_0^T B_T g(\mathbf{p}_i,t) B_T\varphi(\mathbf{p}_i,t) + \lambda \int_0^T \int_\Omega g\varphi \\ \mathcal{L}\varphi &= \sum_{i=1}^{n} B_T\varphi(\mathbf{p}_i,t)(B_T u(\mathbf{p}_i,t) + f_b(\mathbf{p}_i,t) - z_i(t)) \\ c &= \sum_{i=1}^{n} (B_T u(\mathbf{p}_i,t) + f_b(\mathbf{p}_i,t) - z_i(t))^2. \end{aligned}$$

Clearly $\mathcal{A}(g,\varphi)$ is a bilinear form, since B_T, the pointwise evaluation of a function and the integration on an interval are linear operators. Moreover, it is continuous in $\mathcal{G}_T$; indeed, thanks to the embedding $H^2(\Omega) \subset C(\bar{\Omega})$ if $\Omega \subset \mathbb{R}^d$ with $d \leq 3$ and thanks to (3.5) we have that

$$\int_0^T |B_T g(\mathbf{p}_i,t)|^2 \leq \int_0^T \|B_T g(t)\|^2_{C(\bar{\Omega})} \leq C \int_0^T \|B_T g(t)\|^2_{H^2(\Omega)} \leq C \int_0^T \|g(t)\|^2_{L^2(\Omega)}.$$

We thus obtain that $\mathcal{A}(g,\varphi) \leq (C^2+\lambda) \int_0^T \|g(t)\|_{L^2(\Omega)} \int_0^T \|\varphi(t)\|_{L^2(\Omega)}$.
Finally, the operator $\mathcal{A}(g,\varphi)$ is coercive in $\mathcal{G}_T$, since

$$\mathcal{A}(g,g) = \sum_{i=1}^{n} \int_0^T |B_T g(\mathbf{p}_i,t)|^2 + \lambda \int_0^T \|g(t)\|^2_{L^2(\Omega)} \geq \lambda \int_0^T \|g(t)\|^2_{L^2(\Omega)}.$$

Due to the fact that the bilinear form $\mathcal{A}(\cdot,\cdot)$ is continuous and coercive in $\mathcal{G}_T = L^2(0,T;L^2(\Omega))$, that the operator $\mathcal{L}$ is linear and that c is a constant, Theorem 1 states the existence and the uniqueness of $\hat{g} = \operatorname{argmin}_{g\in\mathcal{G}_T} J_{T,g}(g)$. From the bijectivity of $B_T : L^2(0,T;L^2(\Omega)) \to V_{T,0}$ we deduce the existence and uniqueness of $\hat{f} = f_b + B_T(\hat{g}+u) = \operatorname{argmin}_{f\in V_T} J_T(f)$. □

Similarly to Chapter 1, the solution of Problem 5 is obtained by solving:

$$\begin{cases} \frac{\partial \hat{f}}{\partial t} + L\hat{f} = u + \hat{g} & \text{in } \Omega \times (0,T) \\ \hat{f}(\mathbf{x},0) = s(\mathbf{x}) & \text{in } \Omega \\ \mathcal{B}_c \hat{f} = h & \text{on } \partial\Omega \times (0,T] \end{cases}$$
$$\begin{cases} -\frac{\partial \hat{g}}{\partial t} + L^*\hat{g} = -\frac{1}{\lambda}\sum_{i=1}^{n}(\hat{f} - z_i)\delta_{\mathbf{p}_i} & \text{in } \Omega \times (0,T) \\ \hat{g}(\mathbf{x},T) = 0 & \text{in } \Omega \\ {\mathcal{B}_c}^*\hat{g} = 0 & \text{on } \partial\Omega \times (0,T] \end{cases} \tag{3.7}$$

where $\delta_{\mathbf{p}_i}$ is the Dirach mass located in $\mathbf{p}_i$, L^* is the adjoint operator of L defined in (1.5) and $\mathcal{B}_c{}^*$ is the operator that defines the boundary conditions of the adjoint problem. Notice that the dual problem of the parabolic PDE (3.3) is a backward parabolic PDE.

3.4 Finite Element estimator

The estimation Problem 5 cannot be solved analytically and for this reason we approximate the PDE system (3.7) with the Finite Difference method in time and the Finite Element method in space.

Let $\mathcal{T}_h$ be a triangulation of the space domain, where h denotes the characteristic mesh size. We consider the space V_h^r of piecewise continuous polynomial functions of order $r \geq 1$ over the triangulation, defined in (1.10). 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.
For ease of notation, 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 $V_{h,0}^r$, defined in (1.10), which only necessitates of the internal nodes of the triangulation and the associated basis functions, whilst all boundary nodes can be discarded.

In order to define the weak problem associated to the system of PDEs (3.7) we introduce the bilinear form $a(\cdot,\cdot)$ associated to the operator L, defined in (1.12), and the interpolation s_h of the initial value s in V_h^r. The discrete version of the weak problem is thus given by

$$\begin{cases} \int_\Omega \frac{\partial \hat{f}_h}{\partial t}\psi_h + a(\hat{f}_h, \psi_h) - \int_\Omega \hat{g}_h \psi_h = \int_\Omega u \psi_h & t \in (0,T) \\ \hat{f}_h(\mathbf{x}, 0) = s_h(\mathbf{x}) & \\ -\lambda \int_\Omega \frac{\partial \hat{g}_h}{\partial t}\varphi_h + \lambda a(\varphi_h, \hat{g}_h) + \sum_{i=1}^n \hat{f}_h(\mathbf{p}_i, \cdot)\varphi_h(\mathbf{p}_i) = \sum_{i=1}^n z_i \varphi_h(\mathbf{p}_i) & t \in (0,T) \\ \hat{g}_h(\mathbf{x}, T) = 0 & \end{cases} \tag{3.8}$$

for all $\psi_h, \varphi_h \in V_{h,0}^r$, where $\hat{f}_h(\cdot,t), \hat{g}_h(\cdot,t) \in V_{h,0}^r$ $\forall t \in [0,T]$. Notice that the discretization (3.8) of the system of PDEs (3.7) is only a space discretization.

In order to discretize the problem in time we use the Finite Difference method. We consider N_T uniformly spaced points in $[0,T]$, $\tau^0, \ldots, \tau^{N_T}$, such that $\tau^0 = 0$, $\tau^k = k{\cdot}dt$ and $\tau^{N_T} = T$. We define $\hat{f}_h^k(\cdot) = \hat{f}_h(\cdot, \tau^k)$, $\hat{g}_h^k(\cdot) = \hat{g}_h(\cdot, \tau^k)$, $u^k(\cdot) = u(\cdot, \tau^k)$ and $z_i^k = z_i(\tau^k)$ and we discretize the time derivatives as

$$\frac{\hat{f}_h^k - \hat{f}_h^{k-1}}{dt} \approx \frac{\partial \hat{f}_h}{\partial t}.$$

The discretization in time of the system (3.8) is then obtained by means of an Implicit

Euler scheme. The discrete variational system thus becomes:

$$\begin{cases} \int_\Omega \frac{\hat{f}_h^k - \hat{f}_h^{k-1}}{dt}\psi_h + a(\hat{f}_h^k, \psi_h) - \int_\Omega \hat{g}_h^k \psi_h = \int_\Omega u^k \psi_h, \quad k = 1, \dots, N_T \\ \hat{f}_h^0 = s_h \\ -\lambda \int_\Omega \frac{\hat{g}_h^{k+1} - \hat{g}_h^k}{dt}\varphi_h + \lambda a(\varphi_h, \hat{g}_h^k) + \sum_{i=1}^n \hat{f}_h^k(\mathbf{p}_i)\varphi_h(\mathbf{p}_i) \\ \qquad = \sum_{i=1}^n z_i^k \varphi_h(\mathbf{p}_i), \;\; k = 0, \dots, N_T - 1 \\ \hat{g}_h^{N_T} = 0. \end{cases}$$

Defining now $\boldsymbol{\psi} = (\psi_1, \dots, \psi_{N_h})^T$, $\hat{\mathbf{f}}^k$ and $\hat{\mathbf{g}}^k$ respectively the vectors such that $\hat{f}_h^k = \boldsymbol{\psi}^T \hat{\mathbf{f}}^k$ and $\hat{g}_h^k = \boldsymbol{\psi}^T \hat{\mathbf{g}}^k$, we can write this system in matrix form as

$$\begin{cases} (\mathbf{A} + \frac{1}{dt}\mathbf{R})\hat{\mathbf{f}}^k - \frac{1}{dt}\mathbf{R}\hat{\mathbf{f}}^{k-1} - \mathbf{R}\hat{\mathbf{g}}^k = \mathbf{u}^k, & k = 1, \dots, N_T \\ \hat{\mathbf{f}}^0 = \mathbf{s} \\ -\frac{\lambda}{dt}\mathbf{R}\hat{\mathbf{g}}^{k+1} + \lambda(\mathbf{A}^T + \frac{1}{dt}\mathbf{R})\hat{\mathbf{g}}^k + \mathbf{\Psi}^T\mathbf{\Psi}\hat{\mathbf{f}}^k = \mathbf{\Psi}^T\mathbf{z}^k, & k = 0, \dots, N_T - 1 \\ \hat{\mathbf{g}}^{N_T} = 0. \end{cases} \tag{3.9}$$

where the matrices $\mathbf{A}$, $\mathbf{R}$ and $\mathbf{\Psi}$ are defined respectively in (1.17), (1.18) and (1.19), the vectors $\mathbf{z}^k = \left(z_1^k, \dots, z_n^k\right)^T$ contain the value of z_i at time τ^k, $\mathbf{s} = (s_h(\mathbf{p}_1), \dots, s_h(\mathbf{p}_n))^T$, while $\mathbf{u}_j^k = \int_\Omega u^k \psi_j$.

Remark 5. *We can write the system* (3.9) *as a linear system*

$$\begin{bmatrix} \tilde{\mathbf{\Psi}}^T\tilde{\mathbf{\Psi}} & \lambda\tilde{\mathbf{A}}^* \\ \tilde{\mathbf{A}} & -\tilde{\mathbf{R}} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}} \\ \hat{\mathbf{g}} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{\Psi}}^T\mathbf{z} \\ \mathbf{u} \end{bmatrix}$$

where $\hat{\mathbf{f}}^T = ((\hat{\mathbf{f}}^0)^T, \dots, (\hat{\mathbf{f}}^{N_T})^T)$, $\hat{\mathbf{g}}^T = ((\hat{\mathbf{g}}^0)^T, \dots, (\hat{\mathbf{g}}^{N_T})^T)$, $\mathbf{z}$ *contains the observations at different times and the matrices are defined blockwise as functions of the matrices* $\mathbf{A}$, $\mathbf{R}$ *and* $\mathbf{\Psi}$.

3.5 Model for areal data

The time-dependent SR-PDE method can be extended to the case of areal data.
Let $D_i \subset \Omega$, for $i = 1, \dots, N$, be some subdomains and $\bar{z}_i(t)$, for $i = 1, \dots, N$, be the mean value of a quantity of interest on the subdomain D_i at time $t \in [0, T]$. We assume that the observations $\bar{z}_i(t)$ are square integrable random processes with mean

$$\mathbb{E}\left[\bar{z}_i(t)\right] = \frac{1}{|D_i|}\int_{D_i} f_0(\mathbf{p}, t). \tag{3.10}$$

Similarly to the SR-PDE framework for areal data presented in Chapter 1, in order to estimate the surface we propose to minimize the penalized sum-of-square-error functional

$$\bar{J}_T(f) = \sum_{i=1}^{N} \frac{1}{|D_i|} \int_0^T \left(\int_{D_i} (f(\mathbf{x},t) - \bar{z}_i(t))\, d\mathbf{x} \right)^2 dt + \lambda \int_0^T \int_\Omega \left(\frac{\partial f}{\partial t} + Lf - u \right)^2 \tag{3.11}$$

over the space V_T, defined in Section 3.2 in the case of pointwise data. The first term is a data fitting criterion, consisting in a weighted least-square-error functional for areal data over subdomains D_i integrated in time, while the second term is the same roughness penalty term used in the pointwise data framework and previously described.

The estimation problem can thus be formulated as

Problem 6. *Find* $\hat{f} \in V_T$ *such that*

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

Existence and uniqueness of the solution of Problem 6 is provided by the following theorem.

Theorem 7. *Under Assumption 1, the solution of Problem 6 exists and is unique.*

The proof is analogous to the proof of Theorem 6. The existence and uniqueness of the estimator $\hat{f}$ is in fact obtained writing the functional $\bar{J}_T(f)$ as a quadratic form. Notice that we don't need to require that Assumption 4 holds in the areal data framework, since we don't need pointwise evaluations of the dynamic surface.

Similarly to the case of pointwise data, the surface estimator $\hat{f}$ is obtained solving the following system:

$$\begin{cases} \frac{\partial \hat{f}}{\partial t} + L\hat{f} = u + \hat{g} & in\ \Omega \times (0,T) \\ \hat{f}(\mathbf{x},0) = s(\mathbf{x}) & in\ \Omega \\ \mathcal{B}_c \hat{f} = h & on\ \partial\Omega \times (0,T] \end{cases}$$
$$\begin{cases} -\frac{\partial \hat{g}}{\partial t} + L^* \hat{g} = -\frac{1}{\lambda} \sum_{i=1}^{N} \frac{1}{|D_i|} \mathbb{I}_{D_i} \int_{D_i} (\hat{f} - \bar{z}_i) & in\ \Omega \times (0,T) \\ \hat{g}(\mathbf{x},T) = 0 & in\ \Omega \\ \mathcal{B}_c \hat{g} = 0 & on\ \partial\Omega \times (0,T] \end{cases} \tag{3.12}$$

where $\hat{g} \in \mathcal{G}_T$ represents the misfit of the penalized PDE, L^* is the adjoint operator of L, described by equation (1.5), and ${\mathcal{B}_c}^*$ is the operator that defines the boundary conditions of the adjoint problem.

This system of PDEs can be discretized in space by means of the Finite Element method and in time by means of Finite Differences, along the same line followed for pointwise observations in Section 3.4. We can thus write the discrete estimator as the solution of a system analogous to (3.9), where the matrices involved are those defined in Chapter 1 in the areal data framework.

3.6 Simulation study

We test the time-dependent SR-PDE method in a toy example. We consider a circular bidimensional domain $\Omega = B_1$ and we assume that the true underlying surface f_0, represented at time $t = 0$ in the top left panel of Figure 3.2, is defined by

$$f_0(\mathbf{x}, t) = a(t)(1 - \mathbf{x}_1^2 - \mathbf{x}_2^2) \tag{3.13}$$

where $a(t)$ is a smooth function of time, represented in the top right panel of Figure 3.2. We consider the $n = 100$ observation points $\mathbf{p}_1, \dots, \mathbf{p}_n$, represented in Figure 3.2, bottom left.

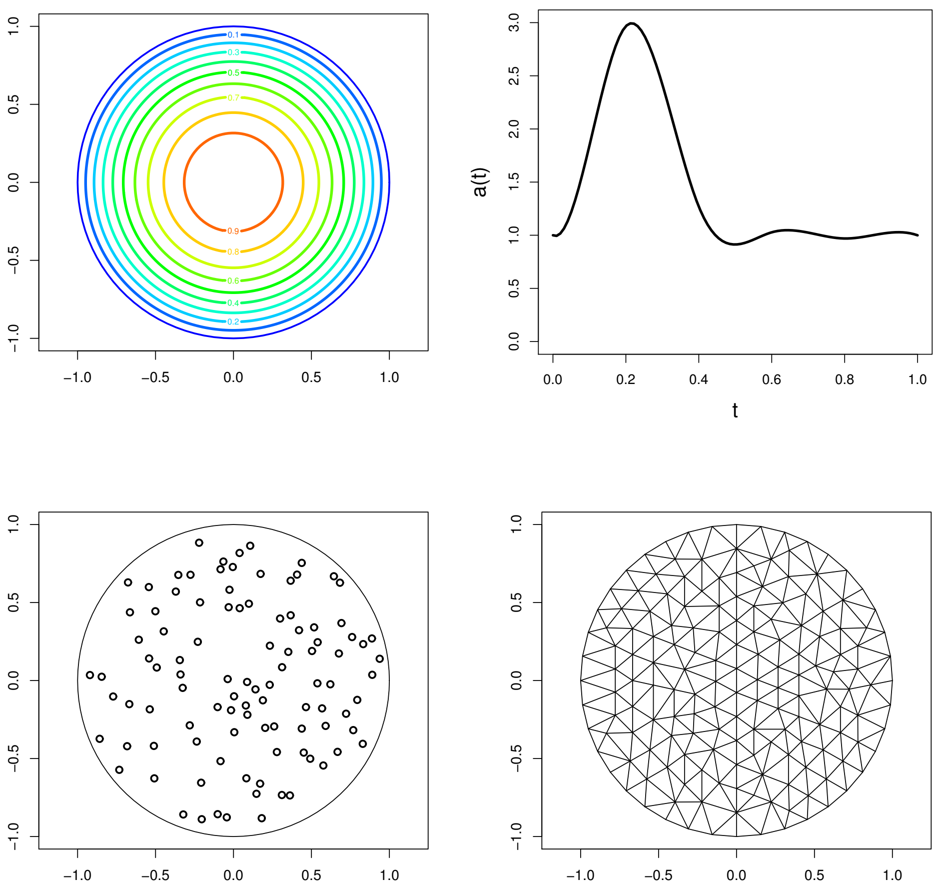

Figure 3.2: *Top left: true surface f_0 at time $t = 0$ used in the toy example; the image displays the isolines $(0, 0.1, \dots, 0.9, 1)$. Top right: time-dependent multiplicative factor $a(t)$ in the true surface f_0. Bottom left: location points $\mathbf{p}_1, \dots, \mathbf{p}_n$. Bottom right: triangulation of the domain Ω.*

We discretize the estimation problem in space using linear Finite Elements defined on the triangulation shown in Figure 3.2, bottom right. The discretization in time is obtained by means of the Finite Difference method, with the implicit scheme described in Section 3.4. The time discretization grid is composed by $N_T = 41$ uniformly spaced points in $[0, 1]$, i.e., $\tau_1 = 0, \tau_2 = 0.025, \dots, \tau_{N_T} = 1$. We add a Gaussian noise with zero mean and standard deviation $\sigma = 0.1$ to the pointwise evaluation $f_0(\mathbf{p}_i, \tau^k)$ of the surface at time τ^k.

The results obtained minimizing the functional $J_T(f)$, penalized with the misfit of the parabolic equation $\frac{\partial f}{\partial t} + \Delta f = 0$ with Dirichlet b.c., $f|_{\partial\Omega} = 0$ and initial condition

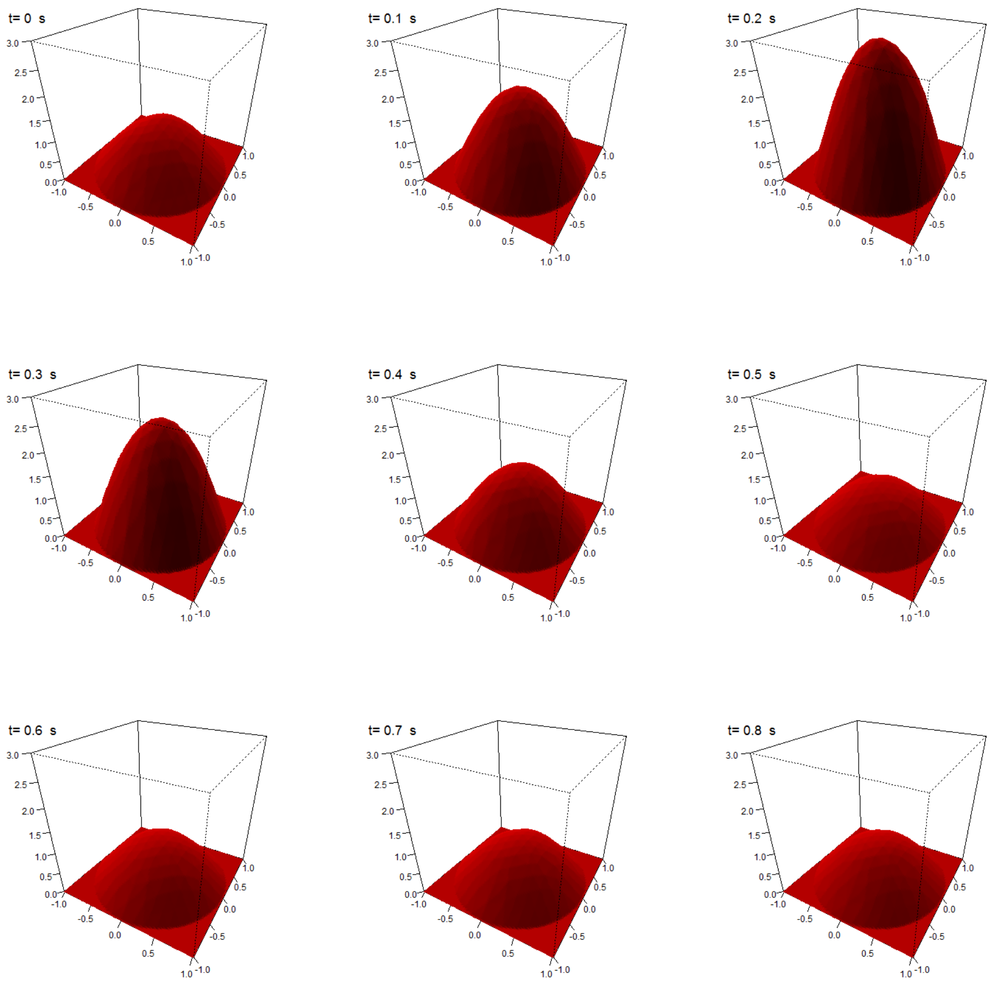

Figure 3.3: *Estimated surface at different time instants using the time-dependent SR-PDE in the toy example.*

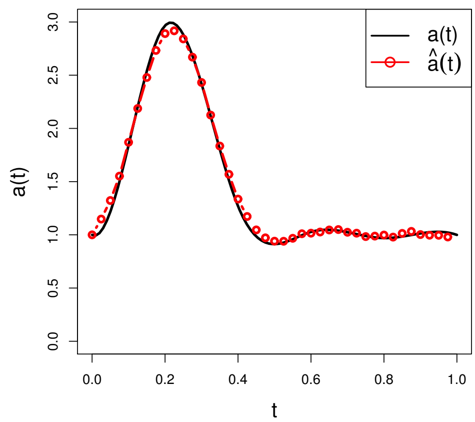

Figure 3.4: *Comparison between the time-dependent multiplicative factor $a(t)$ in the true surface f_0 (black curve) and the maximum value of the estimated surface $\hat{f}_h$, called $\hat{a}(t)$ (red curve).*

$f(\mathbf{x}, 0) = (1 - \mathbf{x}_1^2 - \mathbf{x}_2^2)$, are shown in Figure 3.3. From the comparison between the time-dependent multiplicative factor $a(t)$ in the true surface f_0 and the maximum value of the estimated surface $\hat{f}_h$, represented in Figure 3.4, we can notice that the estimated surface evolving in time captures well the features of the true underlying dynamic surface f_0. Specifically, the estimated surface reaches the maximum peak at time $t = 0.2$ and decreases at time $t = 0.4$ to the plateau level, analogously to the true underlying surface f_0. The maximum value reached by the surface estimator is however lower than the maximum value of f_0 due to the regularizing effect of the method. We can also notice that the method oversmooths the solution in the last time instants; this effect is caused by the fact that, imposing $\hat{g}(\mathbf{x}, T) = 0$, we are actually requiring that the PDE (3.3) is satisfied exactly at time $t = T$. Since we are penalizing in the toy example considered an homogeneous PDE, i.e., $u = 0$, the dynamic surface shrinks to zero in the last time instants.

3.7 Blood velocity field estimation

We want now to apply the time-dependent SR-PDE to the estimation of the dynamic blood velocity field on a cross-section of the common carotid artery. This problem is of particular interest in the MACAREN@MOX project described in Chapter 1. The data used in order to estimate the dynamic blood velocity field are represented in Figure 3.5: the 7 curves represent the mean velocity of the blood cells measured on the 7 beams, whose location is shown in Figure 1.6. Each mean velocity curve is estimated separately from the others starting from the corresponding ECD signal. The period of each signal is estimated using fast Fourier transform and the curve is obtained by means of a Fourier smoothing. The curves are then registered with an affine warping function in order to obtain curves with the same period, equal to the average period $T = 0.92$, and aligned with respect to the systolic peak time, which is an easily detectable landmark. In the upper right part of the figure a stylized artery cross-section is represented and each beam is colored according to the color of the corresponding ECD signal. We

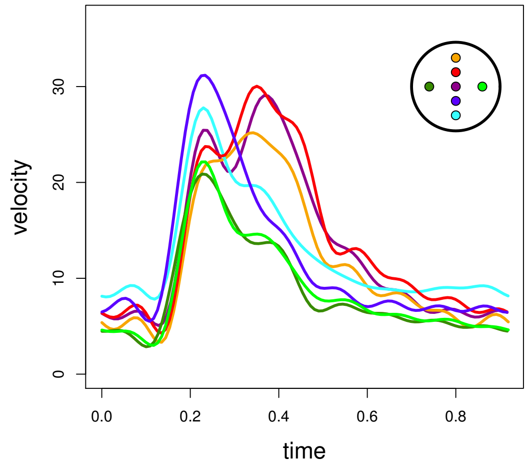

Figure 3.5: *Mean velocity measured on the 7 beams on the artery cross-section. Each curve is colored according to the position of the corresponding beam, represented in the upper right corner.*

can notice that the ECD signals have a different shape. Specifically, the curves corresponding to the central beam and to the beams in the upper part of the section have two peaks in the systolic phase and reach higher velocities during the second peak, while the curves corresponding to the beams in the lower and lateral part of the section have only one peak. In the second group of curves, the ECD signals corresponding to the beams in the lower part of the section reach higher velocities than those corresponding to the lateral beams. The same evolution in time of the mean velocity on the 7 beams is represented in Figure 3.6. This figure represents the observed mean velocity at some fixed time instants. As in Figure 1.6, at each time instant each beam is colored according to the corresponding mean velocity. The data used in Chapter 1 in order to estimate the systolic blood velocity field corresponds to the data shown in the top right panel of Figure 3.6, which represents the mean velocity at the systolic peak time ($t = 0.23$). We can notice that, while at the systolic peak time the maximum velocity is obtained in the lower part of the cross-section, in the following instants the maximum velocity moves to the upper part of the cross-section, as expected from Figure 3.5.

The dynamic velocity field is estimated minimizing the functional $\bar{J}_T(f)$ penalized with the misfit of the parabolic PDE $\dot{f} + \tilde{\lambda} L f = 0$, where the spatial operator L is the same used in Section 1.7 for the estimation of the velocity at the systolic peak time, i.e., the diffusion matrix $\mathbf{K}$ is defined in equation (1.42) with $R = 2.8$, $\kappa_1 = 0.1$, $\kappa_2 = 0.2$, $\mathbf{b}(x, y) = (\beta x, \beta y)^T$ with $\beta = 0.5$. The relative strength between the space and time derivatives is controlled via the multiplying factor $\tilde{\lambda}$, which is set in this case equal to 10 The resulting spatial differential operator has thus diffusion matrix equal to $\tilde{\lambda}\mathbf{K}$ and transport term equal to $\tilde{\lambda}\mathbf{b}$. We estimate the starting velocity profile h_0, which corresponds to the velocity field at the end of the diastolic phase, via the SR-PDE method described in Chapter 1. Finally, we impose homogeneous Dirichlet boundary conditions on the wall of the carotid cross-section, i.e., $f|_{\partial\Omega} = 0$, as in Section 1.7.

The estimation problem is discretized in space using linear Finite Elements defined

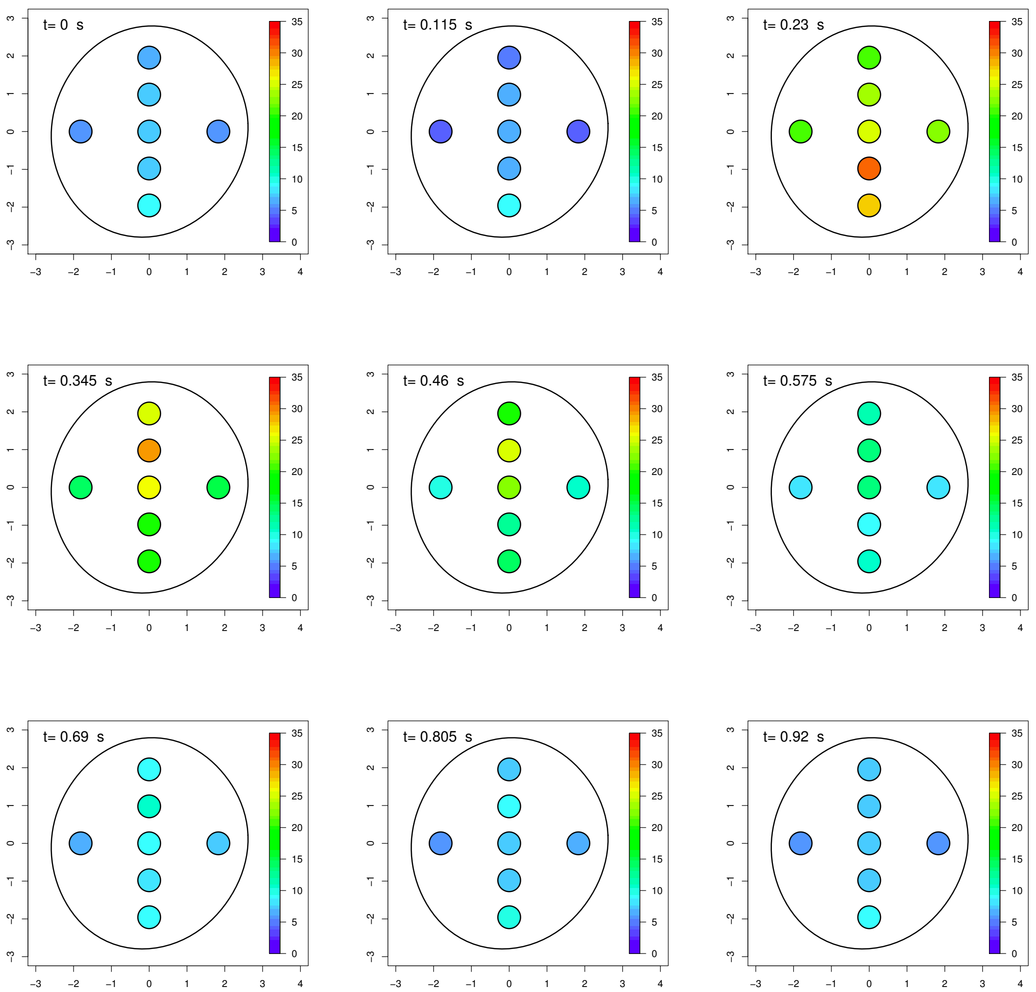

Figure 3.6: *Time evolution of the mean velocity on the 7 beams at certain time instants. Each beam is colored according to the mean blood-velocity measured on the beam at the considered time instant.*

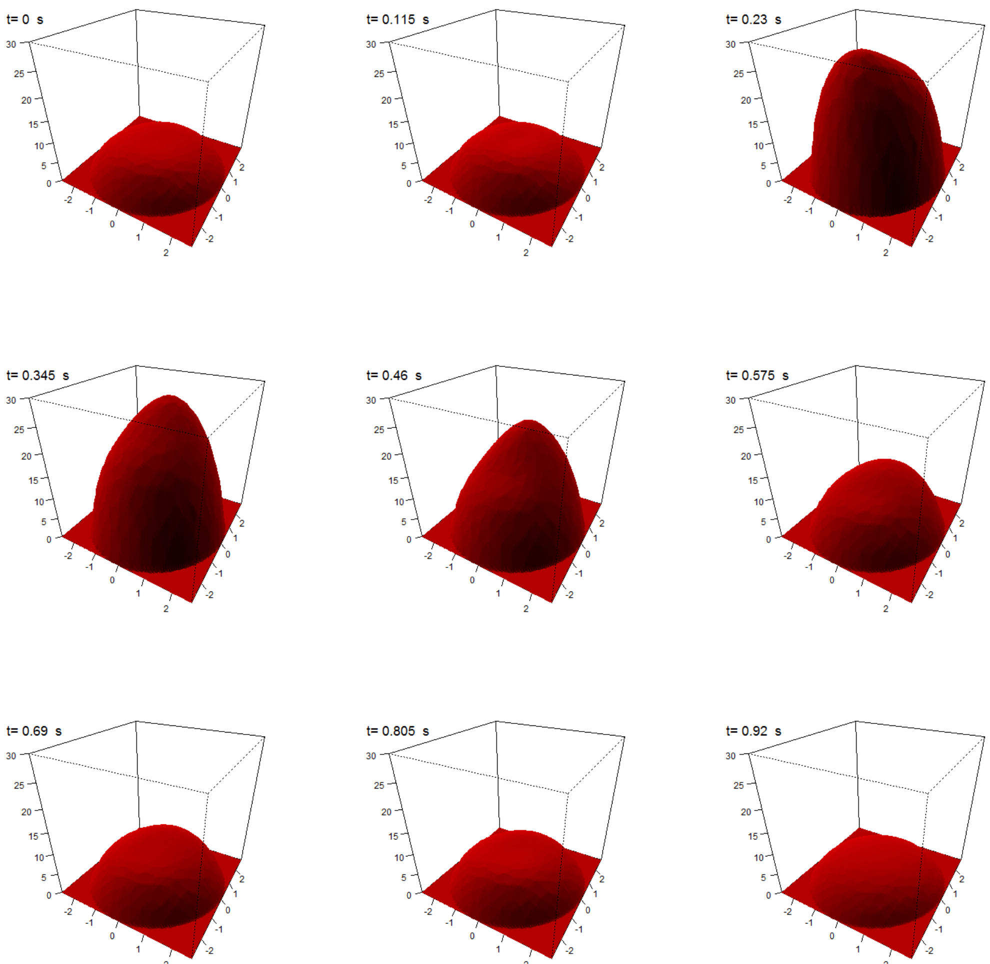

Figure 3.7: *Estimated blood velocity field at different time instants using the time-dependent SR-PDE.*

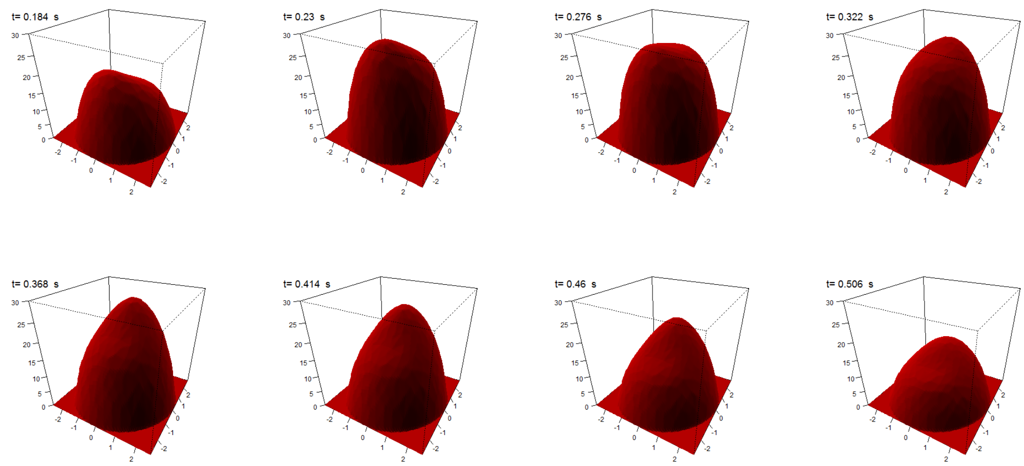

Figure 3.8: *Estimated blood velocity field during the systolic phase using the time-dependent SR-PDE.*

on the triangulation used in the velocity field estimation in Chapter 1, which is shown in the left panel of Figure 1.8. The discretization in time is obtained by means of the Finite Difference method, with the implicit scheme described in Section 3.4. The time discretization grid is composed by $N_T = 41$ uniformly spaced points in $[0, 0.92]$.

The estimated dynamic surface is represented in Figure 3.7 at certain time instants. We can notice that during a heartbeat the shape of the velocity field is subject to strong variations. During the first instants of the systolic phase (upper right panel of Figure 3.7) the velocity field has a strong asymmetry with higher values in the lower part of the artery cross-section. In the following instants the shape of the velocity field changes, assuming higher values in the upper right part of the cross-section (central left and central panel of Figure 3.7). The shape variation of the estimated velocity field during the systolic phase is represented in more detail in Figure 3.8. During the diastolic phase the estimated velocity field is instead symmetric and flat.

In order to assess the goodness of fit we can compare the data, represented in Figure 3.5, with the mean velocity on the beams estimated via SR-PDE. We recall that the data

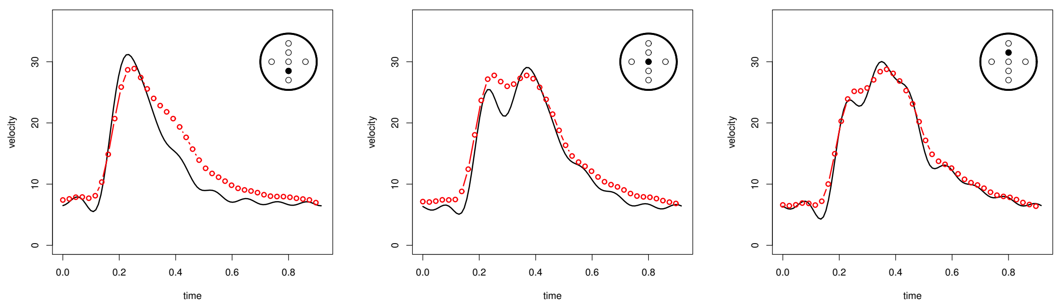

Figure 3.9: *Comparison between the estimated mean velocity (red curves) and the correspondent ECD signals (black curves) on the 3 central beams.*

represent the time-varying mean velocity registered on the beams and that each curve is estimated, separately from the others, starting from the ECD signals. Figure 3.9 shows this comparison on the 3 central beams: the mean velocity estimated via SR-PDE is represented in red, while the correspondent datum is represented in black. Notice that the estimated dynamic surface captures well the main features of the ECD signals. Moreover the estimate of the mean velocity on each beam borrows strength from the proximity of other beams, taking into account the spatial structure of the phenomenon, which is not considered in the estimation of the mean velocity from the ECD signals (black curves in Figure 3.9). Penalizing a parabolic PDE that summarizes the prior knowledge on the phenomenon allows thus to obtain a physiological estimate of the velocity field.

Discussion and future developments

In this thesis 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. The method is actually not applicable to PDEs with discontinuous parameters, pointwise forcing term or defined on irregular domains, due to the extra regularity required to the parameters of the penalized PDE. This request however is not restrictive in spatial statistics and in the smoothing framework since the field is normally assumed to be very regular. The proposed mixed Finite Element method requires moreover g_0 to satisfy the Dirichlet boundary conditions on Γ_D. This hypothesis could be sometimes restrictive since it means that the second derivatives of the field at the boundary are clumped to zero; other discretization methods for fourth order problems could be considered in the future. However, we have observed numerically that whenever g_0 does not satisfy the homogeneous Dirichlet boundary conditions on Γ_D, the extra consistency error is of the same order as the bias contribution and therefore does not really compromise the optimal convergence rate of the method.

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 this thesis, 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 [29].

As shown in Chapter 1, 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. The convergence studied in Chapter 2 concerns the bias of the estimator when the characteristic mesh size h goes to zero and neglects instead the error induced by the presence of noise in the observations. Classical results concerning smoothing splines and thin-plate splines (see, e.g., [7–9, 18]) show the consistency of these estimators when the smoothing parameter λ goes to zero, as $n \to +\infty$, with a proper rate. Unfortunately these results cannot be directly extended to SR-PDE and a different approach needs to be developed to show the consistency of these models. We are currently studying the (infill) asymptotic properties of the estimator when the number of observations n goes to infinity. In particular we are studying the convergence of the variance term $\hat{w}$, both in the continuous and the discrete setting, when n goes to infinity and we are looking for a proper rate of λ that makes both the bias and variance vanish.

It will be also interesting to balance the discretization error induced by the Finite Element approximation with the bias of the estimator and the variance term related to the noise of the observations. A possible way to solve the problem is the development of a proper mesh adaptation technique, based on a posteriori estimates of noise, variance and bias. This technique should locally refine the mesh in order to obtain a local discretization error of the same order or smaller than the bias and the noise standard deviation σ.

Finally, in order to use the estimated dynamic velocity fields as inflow conditions for the hemodynamics simulations, within the MACAREN@MOX project, it will be necessary to extend the time-dependent SR-PDE method in order to account for the deformation of the artery wall during the heartbeat. Figure 3.10 shows the geometry of the artery cross-section located 2 cm before the carotid bifurcation at different time instants during the heartbeat. The black points represent the segmentation of the geometry obtained with Magnetic Resonance Imaging (MRI) data, while the red curves represent the estimated geometry obtained via a space-time periodic smoothing. Notice that during the first instants of the systolic phase, represented in the first panels of Figure 3.10, the carotid cross-section quickly dilates, while after the systolic peak time, represented in the upper right panel of Figure 3.10, the section slowly shrinks. Moreover the shape of the cross-section changes during the whole heartbeat. The domain considered in the blood velocity estimation in the Chapters 1 and 3 corresponds to the carotid cross-section at the systolic peak time that is represented in the upper right panel of Figure 3.10. Allowing for changes in the shape of the domain over time poses a problem of registration of different domains similar to the one faced in population studies, as anticipated in Section 1.7. We are currently developing an appropriate registration method based on shape analysis techniques, landmark registration and conformal mapping to face this issue.

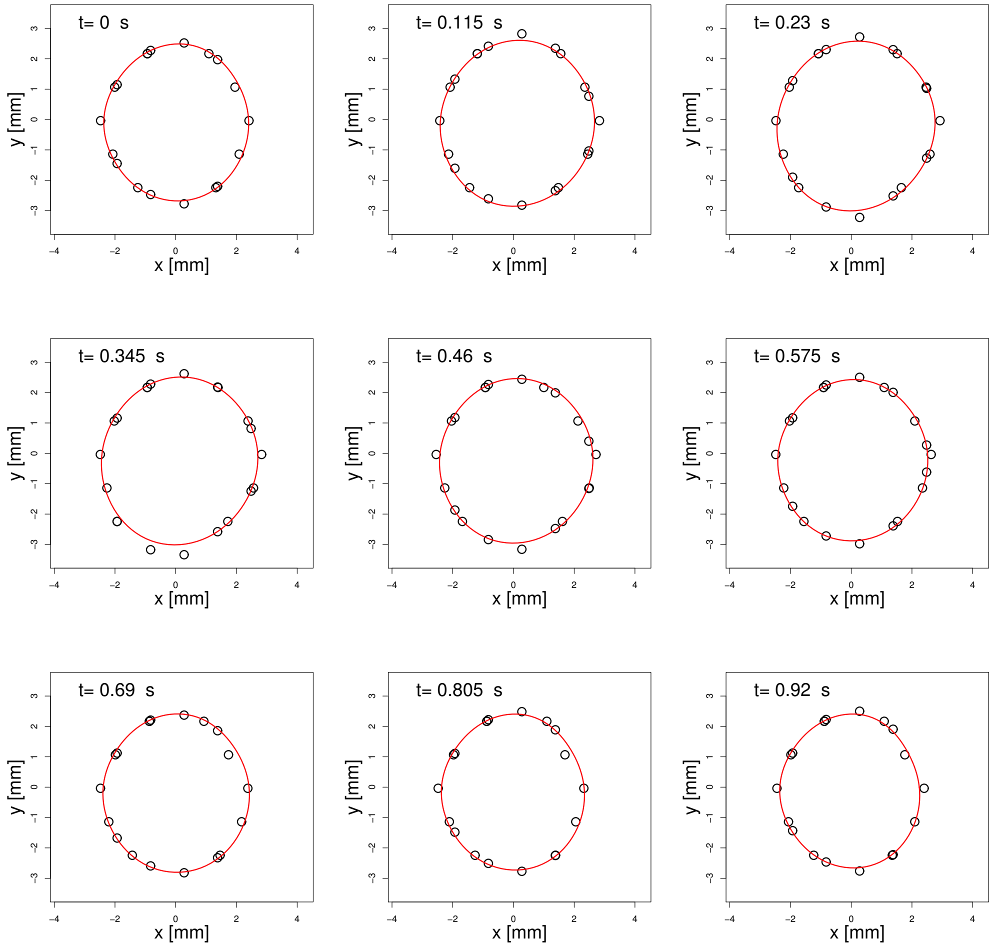

Figure 3.10: *Geometry reconstruction of the carotid cross-section from MRI data during the lapse of a heartbeat: black points represent the geometry obtained from the segmentation of MRI data while red curve represents the estimated geometry obtained via space-time periodic smoothing.*

Bibliography

[1] N.H. Augustin, V.M. Trenkel, S.N. Wood, and P. Lorance. Space-time modelling of blue ling for fisheries stock management. *Environmetrics*, 24(2):109–119, 2013.

[2] S.C. Brenner and L.R. Scott. *The mathematical theory of Finite Element methods*. Springer, 2008.

[3] S.C. Brenner and L.Y. Sung. C^0 interior penalty methods for fourth order elliptic boundary value problems on polygonal domains. *J. Sci. Comput*, 22/23:83–118, 2005.

[4] F. Brezzi and M. Fortin. *Mixed and hybrid finite elements methods*. Springer, 1991.

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

[6] P.G. Ciarlet. *The Finite Element method for elliptic problems*. Society for Industrial and Applied Mathematics, 2002.

[7] D.D. Cox. Asymptotics for m-type smoothing splines. *Annals of Statististics*, 11(2):530–551, 1983.

[8] D.D. Cox. Multivariate smoothing spline functions. *SIAM J. Numer. Anal.*, 21(4):798–813, 1984.

[9] F Cucker and D.X. Zhou. *Learning theory: an approximation theory viewpoint*. Cambridge Monographs on Applied and Computational Mathematics, 2007.

[10] A. Ern and J.L. Guermond. *Theory and Practice of Finite Elements*. Springer, 2004.

[11] 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/.

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

[13] F Ferraty and P. Vieu. *Nonparametric Functional Data Analysis: Theory and Practice*. Springer, 2006.

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

[15] P.L. George and H. Borouchaki. *Delaunay Triangulation and Meshing: Application to Finite Elements*. Butterworth-Heinemann, 1998.

[16] E.H. Georgoulis, P. Houston, and Virtanen J. An a posteriori error indicator for discontinuous Galerkin approximations of fourth-order elliptic problems. *IMA J Numer. Anal.*, (31):281–298, 2011.

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

[18] L. Györfi, M. Kohler, A. Krzyżak, and H. Walk. *A Distribution-Free Theory of Nonparametric Regression*. Springer, 1990.

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

[20] T.J. Hastie, R.J. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer, 2009.

[21] M.J. Lai and L. L. Schumaker. *Spline functions on triangulations*, volume 110 of *Encyclopedia of Mathematics and its Applications*. 2007.

[22] 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.

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

[24] G. Marra, D.L. Miller, and L. Zanin. Modelling the spatiotemporal distribution of the incidence of resident foreign population. *Statistica Neerlandica*, 66(2):133–160, 2012.

[25] 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.

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

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

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

[29] 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.

[30] J.O. Ramsay and B.W. Silverman. *Functional Data Analysis*. Springer, 2005.

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

[32] 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.

[33] G. Stone. *Bivariate splines*. PhD thesis, University of Bath, Bath, 1988.

[34] E. Süli and Mozolevski I. hp-version interior penalty DGFEMs for the biharmonic equation. *Comput. Methods Appl. Mech. Engrg.*, (196):1851 –1863, 2007.

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

[36] G. Wahba. *Spline models for observational data*. Society for Industrial and Applied Mathematics, 1990.

[37] 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.

[38] 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.