Elsevier Editorial System(tm) for Journal Of Computational Physics

Manuscript Draft

Manuscript Number: JCOMP-D-12-00242R2

Title: Optical tomography reconstruction algorithm with the finite element method: An optimal approach with regularization tools

Article Type: Regular Article

Keywords: Inverse problem, Adjoint method, Regularization, Finite element parameterization, Gradient Filtering, Radiative transfer equation

Corresponding Author: Mr. Olivier Balima, Ph.D.

Corresponding Author's Institution: Université du Québec à Chicoutimi

First Author: Olivier Balima, Ph.D.

Order of Authors: Olivier Balima, Ph.D.; Yann Favennec, Ph.D; Daniel Rousse

Abstract: Optical tomography is mathematically treated as a non-linear inverse problem where the optical properties of the probed medium are recovered through the minimization of the errors between the experimental measurements and their predictions with a numerical model at the locations of the detectors. According to the ill-posed behavior of the inverse problem, some regularization tools must be performed and the Tikhonov penalization type is the most commonly used in optical tomography applications. This paper introduces an optimized approach for optical tomography reconstruction with the finite elements methods. An integral form of the cost function is used to take into account the surfaces of the detectors and make the reconstruction compatible with all finite element formulations, continuous and discontinuous. Through a gradient-based algorithm where the adjoint method is used to compute the gradient of the cost function, an alternative inner product is employed for preconditioning the reconstruction algorithm. Moreover, appropriate re-parameterization of the optical properties is performed. These regularization strategies are compared with the classical Tikhonov penalization one. It is shown that both the re-parameterization and the use of the Sobolev cost function gradient are efficient for solving such an ill-posed inverse problem.

Chicoutimi, March 29, 2013

Dr O. Balima
and
Dr. Y. Favennec

to

the reviewers of the paper
JCOMP-D-12-00242

Subject : JCOMP-D-12-00242R1 – "Optical tomography reconstruction algorithm with the finite element method: An optimal approach with regularization tools"

Dear Reviewers,

First of all, we sincerely thank the reviewers who gave us highly interesting suggestions and corrections that enabled us to improve the manuscript quality. The references have been checked with corrections of the typography of "Sobolev".

Below we answer to the second remarks of Reviewers:

> **Question 1:** The parameter l^2 in the definition of the scalar product in Eq.(25) may be related to the size of the the mesh element. As suggested in [1], we agree that the optimization algorithm can be improved with a preconditioned conjugate gradient approaches uisng the Sobolev gradient. Let us point that in recovering the optical properties ($\kappa(x)$ and $\sigma(x)$) a multilinear approach is under study as both coefficient are of different order of magnitude (see O. balima et *al*[2]). Then, ideas suggested in [1] will be taken into account with a sensitivity approach by avoiding line search method that increases the cross-talk in recovering the parameters.
>
> **Question 2:** Yes, others norms can be used and this study was a first approach. Also, the functional norm may be used in future studies.
>
> **Question 3:** Mesh adaptivity near collimated sources will improve only the solution of the forward model. Here, the target is the recovering of the optical properties for medical diagnostic purposes. We currently focus on a constrained adaptivity for the parameters instead of those presented in [3] where the mesh adaptivity reduces the size of the inclusions.

Looking forward to being published in the *Journal of Computational Physics*. Sincerely yours,

O. Balima & Y. Favennec

References

[1] P. Frandsen, K. Nielsen, O. Tingleff, Unconstrained optimization, Technical University of Denmark, 3rd Edition (March 2004).

[2] O. Balima, J. Boulanger, A. Charette, D. Marceau, New developments in frequency domain optical tomography. Part I. forward model and gradient computation, Journal of Quantitative Spectroscopy and Radiative Transfer 112 (7) (2011) 1229–1234.

[3] H. Jiang, Difuse optical tomography : Principles and Applications, Taylor and Francis, 2011.

Novelty

- New strategies to improve the accuracy of the reconstruction through mesh and finite elements parameterization.
- Use of gradient filtering through an alternative inner product within the adjoint method.
- An integral form of the cost function is used to make the reconstruction compatible with all finite element formulations, continuous and discontinuous.
- Gradient-based algorithm with the adjoint method is used for the reconstruction.

Optical tomography reconstruction algorithm with the finite element method: An optimal approach with regularization tools

O. Balima[a,1,*], Y. Favennec[c], D. Rousse[a]

[a] *Département des Sciences Appliquées, Université du Québec à Chicoutimi, 555 bd de l'Université, Chicoutimi, QC, G7H 2B1, Canada.*
[b] *Chaire de recherche industrielle en technologies de l'énergie et en efficacité énergétique (t3e), École de technologie supérieure 201 Boul. Mgr, Bourget Lévis (QC) Canada, G6V 6Z3.*
[c] *LTN UMR CNRS 6607 - Polytech' Nantes - La Chantrerie, Rue Christian Pauc, BP 50609 44 306 Nantes Cedex 3, France.*

Abstract

Optical tomography is mathematically treated as a non-linear inverse problem where the optical properties of the probed medium are recovered through the minimization of the errors between the experimental measurements and their predictions with a numerical model at the locations of the detectors. According to the ill-posed behavior of the inverse problem, some regularization tools must be performed and the Tikhonov penalization type is the most commonly used in optical tomography applications. This paper introduces an optimized approach for optical tomography reconstruction with the finite elements methods. An integral form of the cost function is used to take into account the surfaces of the detectors and make the reconstruction compatible with all finite element formulations, continuous and discontinuous. Through a gradient-based algorithm where the adjoint method is used to compute the gradient of the cost function, an alternative inner product is employed for preconditioning the reconstruction algorithm. Moreover, appropriate re-parameterization of the optical properties is performed. These regularization strategies are compared with the classical Tikhonov penalization one. It is shown that both the re-parameterization and the use of the Sobolev cost function gradient are efficient for solving such an ill-posed inverse problem.

Keywords: Inverse problem, Adjoint method, Regularization, Finite element parameterization, Gradient Filtering, Radiative transfer equation

1. Introduction

Among the new imaging modalities expected to be available in the future, optical tomography is one of the most promising. It is used in flow diagnostics, medical imaging, food processing, etc. This laser-based probing technique may be divided into direct imaging where the emerging signal is directly used for

*Corresponding author
Email address: ofbalima@gmail.com (O. Balima)

projection and reconstruction imaging based on the solution of an inverse problem. For both of them, recent research tends to show that the use of the long term photons, which have travelled for a long time in the whole sample to be probed, generates more information to the image reconstruction [1].

In direct tomography, a measurable variable of the transmitted or / and reflected signals is processed in order to extract some information about the inside of a semi-transparent medium on which a laser beam has been applied. In material with a high level of scattering, direct tomography is of limited use because photons do not progress along a straight line and the reconstruction is therefore non-linear, which prohibits the use of direct reconstruction methods such as the Radon transform method [2]. The other method, also called optical tomography, is an inverse-based reconstruction technique where the optical properties are estimated from boundary measurements of transmitted light, providing a non-invasive diagnostic tool for medical applications as the optical properties are related to the pathological or physiological state of tissues. The reconstruction is done by minimizing a cost function that measures the errors between the experimental measurements at the detectors and their prediction with a numerical model [1, 3].

Major improvements have been carried out in the last decades using the full radiative transfer equation as the forward model as this equation well describes light transport in biological tissues. Then, different forms of the RTE have been used [1, 4, 5, 6] among which the frequency domain formulation is the most used. The frequency domain approach provides an alternative to scientists to avoid the technical limitations intrinsic to the use of the time-domain approach. It allows a better separation of the optical properties by giving some additional information (phase shift) compared to the stationary domain [7]. Also, highly accurate numerical formulations of the RTE have been achieved such as the discontinuous Galerkin finite element formulation [8, 9, 10]. This method uses numerical fluxes to achieve local conservativity [11, 12, 13] compared to continuous finite element formulations that suffer from the lack of local conservativity.

The ill-posed behavior of the inverse problem requires the use of accurate forward model of light transport coupled with robust optimization techniques. Also, the noise which is inherent by nature to the measurements leads to a non-smooth gradient of the cost function when the classical L_2 inner product is used within the adjoint method and thus the related optimization procedure may be slow and less accurate. In addition, parameters can be of different orders of magnitude, which often leads to problems of crost-talk. Generally, the method of Tikhonov regularization is used to reduce these difficulties. This technique is based on a penalization of the difference between the obtained properties and some guessed ones chosen a priori (usually the background) within the cost function to be minimized. This old regularization method [14] which provides information to the inverse problem actually stabilizes it. However, the choice of the weigh related to the penalization term is problematic since it often leans on the search of a particular region on the so-called L-curve [15]. The use of alternative inner products when extracting the gradient of the cost function, as initiated by [16], aims at smoothing the gradient and acts as a preconditionner for the reconstruction optimization problem. Also, a re-parameterization of the functional space related to the optical properties

is performed in order to avoid over-parameterization with respect to the lack of measurement information.

Then, this paper focuses on new strategies to improve the reconstruction in optical tomography by using the Sobolev gradient (gradient filtering) with finite element parameterization (mesh and space approximation) of the optical parameters. For this purposes, an optimization of the reconstruction scheme is introduced through the choice of an inner product within the adjoint method for the computation of the cost function gradient. The adjoint equations is derived from the continuous radiative transfer equations (CRTE). This leads, when choosing different discretization schemes, rather than reconsidering the adjoint equations, to choose a numerical scheme for the adjoint problem that is coherent with the one chosen for the forward problem. We thus chose the so-called "Differentiate-then-Discretize" approach as opposed to the "Discretize-then-Differentiate" approach as defined in [17] for the simplicity and conciseness when deriving the adjoint equations, and especially because we make use of different functional spaces (different meshes) for the states, the adjoint states, the optical properties, etc.

The paper is organized as follows. Section 2 presents the forward model equations describing the radiative transfer equation along with the measurable quantity used for optical tomography purposes. Section 3 states the cost function that is to be minimized, writes down the optimization problem and describes carefully the adjoint problem based on the continuous radiative transfer equation along with the cost function directional derivatives. Section 4 gives some specific tools that are to be used to cope with the ill-posed nature of the inverse problem, *i.e.* the use of regularization. Specially, the classical Tikhonov regularization is presented with its pros and cons. Other strategies such as the use of an appropriate finite element (mesh and space) parameterization of the optical properties and the use of the Sobolev gradient instead of the usual Hilbert one are of interest. Numerical tests are performed on the presented regularization tools with a comparative analysis. Especially, the use of the Sobolev Gradient and the appropriate re-parameterization is compared with the classical use of the Tikhonov regularization. The last section deals with the conclusions and extended future work.

2. Forward model equations

2.1. Model equations

In optical tomography, the forward model is a numerical model of light transport within the tissues. It aims at computing the prediction of the measurements at the detectors once the source and the optical properties of the medium are known. This model is described by a Boltzmann type integro-differential equation called the radiative transfer equation [1, 18]. This equation is difficult to solve and analytical solutions are available only for simple cases. Below we present the equations of the forward model and of the measurement prediction.

The forward model used is this study is the frequency domain form of the radiative transfer equation which writes [3]:

$$\left(\vec{\Omega}\cdot\nabla+\frac{i\omega}{c}+\kappa+\sigma\right)I(r,\vec{\Omega},\omega)=\frac{\sigma}{4\pi}\int_{4\pi}I(r,\vec{\Omega}^{'},\omega)\Phi(\vec{\Omega}^{'},\vec{\Omega})\mathrm{d}\Omega^{'} \tag{1}$$

where $r\in\mathcal{D}$ is the spatial position, $\vec{\Omega}$ is the propagation direction of light, $i=\sqrt{-1}$, c is the light speed in the medium, ω is the modulation frequency, $I(r,\vec{\Omega},\omega)$ is the radiant power per unit solid angle per unit area at spatial position r in direction $\vec{\Omega}$, $\kappa=\kappa(r)$ and $\sigma=\sigma(r)$ are respectively the absorption and the scattering coefficients. $\Phi(\vec{\Omega}^{'},\vec{\Omega})$ is the scattering phase function. Usually, the scattering phase function in tissues is described by the Henyey-Greenstein phase function [19]:

$$\Phi(\vec{\Omega}^{'},\vec{\Omega})=\frac{1-g^2}{\left(1+g^2-2g\cos(\phi)\right)^{3/2}} \tag{2}$$

where ϕ is the scattering angle between directions $\vec{\Omega}$ and $\vec{\Omega}^{'}$ and g is the anisotropy factor *i.e* the mean cosine. The boundary condition for Eq (1) in the absence of reflecting boundary condition, is an external collimated beam that penetrates into the medium at the spatial position r_0, with direction $\vec{\Omega}_c(r_0)$ such that:

$$I(r,\vec{\Omega},\omega)=q_0(r,\omega)\mathbb{1}_{[r\in\partial\mathcal{D}_0]}\delta(\vec{\Omega}-\vec{\Omega}_c(r_0)) \tag{3}$$

for $\vec{\Omega}\cdot\vec{n}<0$ and $\forall r\in\partial\mathcal{D}$; where q_0 is the collimated light flux density, δ is the Dirac delta function, $\vec{n}$ is the outward unit vector at the surface, and $\mathbb{1}_{[\,\cdot\,]}$ equals one or zero depending on whether or not $r\in\partial\mathcal{D}_0$.

In order to handle strong discontinuities due to the boundary conditions of Dirichlet type that are used in optical tomography with collimated beams, the radiative intensity is separated into two components $I=I_c+I_s$, respectively the collimated (I_c) and scattered intensity (I_s) within the medium [19, 20]. The collimated component obeys the extinction law [19]

$$\left(\frac{i\omega}{c}+\vec{\Omega}_c\cdot\nabla+(\kappa+\sigma)\right)I_c\left(r,\omega\right)=0 \tag{4}$$

where the boundary condition is given by

$$I_c\left(r,\omega\right)=q_0(r,\omega)\mathbb{1}_{[r\in\partial\mathcal{D}_0]},\quad\vec{\Omega}_c\cdot\vec{n}<0\quad\forall r\in\partial\mathcal{D}, \tag{5}$$

whereas the scattered component I_s is the solution of

$$\left(\vec{\Omega}\cdot\nabla+\frac{i\omega}{c}+\kappa+\sigma\right)I_s(r,\vec{\Omega},\omega)=\frac{\sigma}{4\pi}\int_{4\pi}I_s(r,\vec{\Omega}^{'},\omega)\Phi\left(\vec{\Omega}^{'},\vec{\Omega}\right)\mathrm{d}\Omega^{'}+S_c \tag{6}$$

where $S_c=S_c(r,\vec{\Omega},\omega)$ is the source term induced by the scattering of the collimated (I_c) within the medium. The source S_c is given by

$$S_c=\frac{\sigma}{4\pi}\int_{4\pi}I_c(r,\omega)\delta(\vec{\Omega}^{'}-\vec{\Omega}_c)\Phi(\vec{\Omega}^{'},\vec{\Omega})\mathrm{d}\Omega^{'} \tag{7}$$

Let us recall that Eq (6) uses only a vacuum boundary condition, *i.e.*

$$I_s(r,\vec{\Omega},\omega)=0,\quad\vec{\Omega}\cdot\vec{n}<0,\quad\forall r\in\partial\mathcal{D}. \tag{8}$$

Then, the forward model equations are given by Eq.(4) and Eq.(6) associated to their corresponding boundary conditions, Eq.(5) and Eq.(8) respectively. Once, the solutions I_c and I_s are known, the measurable quantity for optical tomography purposes is given by

$$P(I_s) = \int_{\vec{\Omega}\cdot\vec{n}>0} I_s(r,\vec{\Omega},\omega)\vec{\Omega}\cdot\vec{n}\,\mathrm{d}\Omega \quad \forall r \in \partial\mathcal{D} \tag{9}$$

which is the emerging intensity from the medium where $\partial\mathcal{D}$ is the boundary of the domain.

In Eq. (9), it is generally assumed that the contribution of the collimated part of the intensity to the measurements is negligible because of the large extinction of the collimated intensity within the medium. Information is obtained through the induced scattered part of the intensity [21]. However, the collimated equation will be used in the computation of the adjoint equation and the gradient of the objective function as shown in [22].

2.2. Solution method of the forward model

In order to compute the boundary measurement by Eq. (9), one has to solve numerically the forward model equations (Eq. (4) and Eq. (6)) before computing the measurable quantity $P(I_s)$. These equations are difficult to solve and analytical solutions are available only for simple cases or with some strong approximations. Then, numerical solution methods are used among which the finite volume and finite difference schemes associated to the discrete ordinates methods are the most used. These numerical methods are well suited to advection type equations due to their conservative properties, but are less flexible than the finite elements methods in handling complex geometries.

Recently in optical tomography application with the full radiative transfer equation, an increasing interest has been devoted to the finite elements formulations due to its simplicity, flexibility and property of being able to handle complex geometries and advection type equations.

In this study, the solution method is based on the Discontinuous finite elements formulation associated to the discrete ordinates methods. This finite element method approach makes use of numerical fluxes to achieve local conservativity [12, 13] while keeping the flexibility and the property of handling easily complex geometries. In addition, with the same formulation, the finite volume method is obtained when constant polynomial elements are used , while using a continuous function yields the standard finite element method [11]. We refer to [10, 20] for all details on the numerical procedure and its accuracy. Let us remark that the same numerical procedure will be applied to solve the adjoint problem as both, the forward and the adjoint problems are of the same type of advection equations.

3. Inverse model equations

Optical tomography is an inverse problem which aims at recovering the optical properties of the medium under investigation. This is done through the minimization of a cost function that is based on the errors

between the measurements and the predictions of the forward model.

We consider below an integral cost function and present the gradient expressions with the reconstruction algorithm. Different strategies for improving the reconstruction are presented with the use of appropriate parameterization of the parameter space, the use of Tikhonov type regularization and the use of filtering of the gradient through an optical choice of the inner product that is involved in the adjoint method.

3.1. Cost function

The cost function used in optical tomography is generally expressed as a discrete sum of the square of the errors between the measurements and the predictions of the model at the locations of the detectors [1, 22, 23, 24]. However, in practical applications, the detectors cannot be reduced topologically to a probing point and are better described by a small surface at the boundary which may correspond to several nodes of the boundary mesh in case of the finite element discretization. This leads to the impracticability of some finite element formulations like the discontinuous Galerkin method where the basis functions are not well defined on the boundary of the mesh elements [11], whereas their integrals are. The integral form of the cost function makes possible to take into account the surface of the detectors in optical tomography applications while keeping compatibility with all finite elements formulations [25]. The integral cost function writes

$$\mathcal{J}\left(I_s\right) = \frac{1}{2}\sum\!\!\!\!\!\!\int \|P(r,\omega,I_s) - M(r_d,\omega_s)\|^2 \tag{10}$$

where the symbol $\sum\!\!\!\!\!\!\int[\,\cdot\,]$ represents

$$\sum_{s=1}^{N_s}\sum_{d=1}^{N_d}\int_{\partial\mathcal{D}_d} \cdot \ \mathrm{d}s \tag{11}$$

and $\partial\mathcal{D}_d$ is the surface of the d^{th} detector such that $\partial\mathcal{D}_d \subset \partial\mathcal{D}$, $\partial\mathcal{D}$ being the boundary of the domain whose properties are to be recovered. N_s is the number of collimated sources, N_d is the number of detectors, $P(r,\omega,I_s)$ is the prediction of the measurements $M(r_d,\omega_s)$ at the d^{th} detector located at the position r_d for the s^{th} source at the modulation frequency ω_s, and $\|.\|^2$ is defined by

$$\|z\|^2 = z\bar{z} \quad \forall z \in \mathbb{C} \tag{12}$$

where $\bar{z}_i$ is the conjugate of z_i. Let us remark that the norms used to build both, the discrete cost function and the integral one are equivalent mathematically, however the integral cost function has a numerical filtering effect [25] which improves the reconstruction. Moreover, the integral form for the cost function is coherent with practical applications.

The cost function to be minimized may incorporate some penalizing terms that limit the magnitude of the parameters to be retrieved. To do so, let κ and σ belong to the same control space, say for instance $L_2(\mathcal{D})$, set $\theta = (\kappa,\sigma)$, $\theta \in \mathcal{U} = L_2(\mathcal{D}) \times L_2(\mathcal{D})$ and add to the cost function (10) the Tikhonov penalization

function [26], where θ_0 may be the guess:

$$\mathcal{J}^+(\theta) = \frac{\epsilon}{2} \|\theta - \theta_0\|^2_{\mathcal{U}} \tag{13}$$

3.2. Optimization problem

The inverse problem consists in solving the following minimization problem

$$\inf_{\theta \in \mathcal{U}} j(\theta) \tag{14}$$

where $j(\theta)$ equals by definition to the penalized cost function, *i.e.*

$$j(\theta) := \mathcal{J}(I_s) + \mathcal{J}^+(\theta) \tag{15}$$

It is assumed that the cost function $j(\theta)$ is sufficiently smooth and that it is differentiable with respect to θ in order to use gradient algorithms for solving the optimization problem (14). The way the gradient of the cost function is derived is given afterwards in subsection 3.3. The knowledge of the gradient of the cost function for a point location θ enables the use of efficient gradient-type optimization algorithms such as the L-BFGS [27, 28, 29, 30] associated to an inexact line search [23] recalled briefly in subsection 3.4.

The optical tomography problem being ill-posed, specific strategies should be used to improve the reconstruction. Added to the regularization through the Tikhonov penalization, regularization considering a better parameterization of the control space and the use of alternative inner products when extracting the cost function gradient are to be considered. Indded, alternative inner products such as the $H^1(\mathcal{D})$ inner product, or derivative of it, rather than the classical $L_2(\mathcal{D})$ inner product acts as a preconditioner within the optimization procedure while filtering the noise that appears in the gradient of the cost function. These approaches are detailed further in Section 4.

3.3. Gradient computation with the adjoint method

The computation of the cost function gradient may be time consuming in optical tomography applications. Indeed, the related problem is nonlinear and very large when the full radiative transfer equation is used as the forward model. Then, the adjoint method gives an efficient way to compute the gradient by solving an additional equation for the adjoint variable whose cost is equivalent to that of the forward problem and then, evaluating the gradient through a simple inner product [22, 31, 32, 33, 34, 35].

Appendix A gives the methodology for writing down the adjoint equations and gradient components when minimizing the cost function $j(\theta)$, the forward model equations being taken as implicit constraints.

Following Appendix A, the coupled adjoint problems are written as:

$$\begin{cases} \left[-\vec{\Omega}\cdot\nabla - \dfrac{i\omega}{c} + \kappa + \sigma_s\right] I_s^*(r,\omega) - \dfrac{\sigma}{4\pi}\displaystyle\int_{4\pi} I_s^*\Phi(\vec{\Omega}',\vec{\Omega})\,\mathrm{d}\Omega' = 0 & \forall x \in \mathcal{D} \\ \vec{\Omega}\cdot\vec{n}I_s^* + (P-M)\,P' = 0 & \forall x \in \partial\mathcal{D}^- \\ \left[-\vec{\Omega}\cdot\nabla - \dfrac{i\omega}{c} + \kappa + \sigma_s\right] I_c^*(r,\omega) - \dfrac{\sigma}{4\pi}\displaystyle\int_{4\pi} I_s^*\delta(\vec{\Omega}'-\vec{\Omega}_c)\Phi(\vec{\Omega}',\vec{\Omega})\,\mathrm{d}\Omega' = 0 & \forall x \in \mathcal{D} \\ I_c^* = 0 & \forall x \in \partial\mathcal{D}^- \end{cases} \tag{16}$$

and the components of the gradient of the objective function are given by

$$j'(\theta;\kappa') = \langle \kappa' I_c, I_c^*\rangle_{L_2(\mathcal{D})} + \langle \kappa' I_s, I_s^*\rangle_{L_2(\mathcal{D})} + \{\epsilon(\kappa-\kappa_0),\kappa'\}_{\mathcal{U}} \tag{17}$$

$$\begin{aligned} j'(\theta;\sigma') &= \langle \sigma' I_c, I_c^*\rangle_{L_2(\mathcal{D})} + \langle \sigma' I_s, I_s^*\rangle_{L_2(\mathcal{D})} + \{\epsilon(\sigma-\sigma_0),\sigma'\}_{\mathcal{U}} \\ &- \left\langle \frac{\sigma'}{4\pi}\int_{4\pi}\left[I_s(r,\vec{\Omega}',\omega) + I_c(r,\omega)\delta(\vec{\Omega}'-\vec{\Omega}_c)\right]\Phi(\vec{\Omega}',\vec{\Omega})\,\mathrm{d}\Omega', I_s^*\right\rangle_{L_2(\mathcal{D})} \end{aligned} \tag{18}$$

with null boundary conditions for both I_s^* and I_c^* as the boundary condition of the forward model does not depend on any optical parameter.

The inner products involved in $(\nabla j(\theta),\theta') = j'(\theta;\theta')$ used to extract the cost function gradient from (A.16)-(A.17) is to be defined. The most usual definitions for this inner product is the $L_2(\mathcal{D})$ product, that is $(u,v) := \langle u,v\rangle_{L_2(\mathcal{D})}$ as defined in (A.10), but some alternative norms may also be considered as will be presented in next section. With such inner products definition, i.e. $j'(\theta;\theta') = \langle \nabla j,\theta'\rangle_{L_2(\mathcal{D})}$, one obtains straight-forwardly:

$$\nabla_\kappa j(\theta) = \overline{I_c} I_c^* + \overline{I_s} I_s^* + \epsilon(\kappa-\kappa_0) \tag{19}$$

$$\begin{aligned} \nabla_\sigma j(\theta) &= \overline{I_c} I_c^* + \overline{I_s} I_s^* + \epsilon(\sigma-\sigma_0) \\ &- \frac{\overline{I_s^*}}{4\pi}\int_{4\pi}\left[I_s(r,\vec{\Omega}',\omega) + I_c(r,\omega)\delta(\vec{\Omega}'-\vec{\Omega}_c)\right]\Phi(\vec{\Omega}',\vec{\Omega})\,\mathrm{d}\Omega' \end{aligned} \tag{20}$$

Such an inner product definition leads to consider, when the adjoint states are known, the cost function gradients simply with a sum of products (see the right hand side of (A.18)-(A.19)) where the adjoint states are given through the integration of the coupled system (A.15).

3.4. Optimization algorithm

The reconstruction scheme is based on the Limited memory Broyden-Fletcher-Goldfarb-Shanno method (L-BFGS) method which has been efficient for large scaled optimization problems particularly in optical tomography due to the spatial dependency of the optical parameters. The iterative procedure is as follows:

$$\theta_{k+1} = \theta_k + \alpha_k d_k \tag{21}$$

where

$$d_{k+1} = -\nabla j(\theta) + \gamma s_k + \mu y_k \tag{22}$$

where α_k is the step length, d_k is the descent direction. Vectors s_k and y_k are defined by

$$s_k = d_{k+1} - d_k, \quad y_k = \nabla j(\theta_{k+1}) - \nabla j(\theta_k) \tag{23}$$

and coefficients γ and μ by

$$\begin{aligned} \gamma &= -\left(1 + \frac{y_k^T y_k}{s_k^T y_k}\right) \frac{s_k^T \nabla j(\theta_{k+1})}{s_k^T y_k} + \frac{y_k^T \nabla j(\theta_{k+1})}{s_k^T y_k} \\ \mu &= \frac{s_k^T \nabla j(\theta_{k+1})}{s_k^T y_k} \end{aligned} \tag{24}$$

This iterative procedure is associate to an inexact line search of Armijo type for the evaluation of step length α_k since the evaluation of the objective function is time consuming due to the use of the full radiative transfer equation that is difficult to solve. Scaling of both objective functions with the measurements and their gradient independently are done in order to handle round-off errors as in [23] where the efficiency of the algorithm is shown.

4. Specific tools due to the optical tomography ill-posed nature

The inverse problem of recovering the optical properties in optical tomography may be severely ill-posed because on one hand, the data contains noised measurements and especially because the amount of data is not sufficient to retrieve the large amount of unknowns. Hence some regularization strategies must be performed in order to provide additional information (subsection 4.1), reduce the number of unknowns as the measurements are limited (subsections 4.2.1 and 4.2.2), impose regularity constraints with an adapted parameterization of the space of the unknowns (subsection 4.3), and filter the noise that is inherent to measurements, not on the measurements but rather on the cost function gradient itself (subsection 4.3).

4.1. Tikhonov regularization

This is the first introduced regularization approach [14, 36, 26, 37] which consists in providing information to the system in order to stabilize it. Recently, the Gauss–Newton method and its modified versions such as the Levenberg–Marquardt have been extensively used these last years, see e.g. [24, 38, 39]. Though these methods converge almost quadratically, a linear system has to be solved at each iteration of the optimization procedure. The matrix involved in the Gauss–Newton procedure is likely to be badly conditioned, and a damping is thus to be performed, for instance though the use of the Tikhonov penalization term $\epsilon \|\theta - \theta_0\|^2$ involved in (13) and its derivative version in following equations. Such regularization may also be used together with gradient-conjugate type methods and other quasi-Newton methods such as the DFP or the

BFGS approaches [28, 30, 40]. However, the effect of the regularization is a bit attenuated on these latest methods because there is no linear system to be solved since the inverse hessian itself is updated at each iteration, thus the recovering of the properties is a bit less sensitive to the ill-posed nature of the inverse problem.

Moreover, the choice for the regularization parameter ϵ remains problematic particularly for nonlinear inverse problem. The most popular method to determine an appropriate value for ϵ leans on the L-curve method as proposed by Hansen [15]. This method leans on a compromise between the residuals $\mathcal{J}$ and $\mathcal{J}^+$ while such curve parameterized in ϵ presents a positive curvature region where the parameter has to be chosen on. Though this method is often efficient, it is sometimes uneasy to use when the curvature is low and the method is time consuming since one full reconstruction must be realized for each value of ϵ in order to plot such L-curve.

4.2. Regularization by parameterization

An other ways of reducing the ill-posed behavior of the inverse problem is the parameterization of the space of the unknowns as stated in [41]. This approach is of capital importance as it can reduce and simplify the ill-posed behavior of the inverse problem. We refer to [42] where a re-parameterization of the parameters (named reduced model in the corresponding study) has allowed an improvement of the estimated reduced model. Here, these ideas are applied in the finite element framework to parametrize (mesh and finite element space) the unknown parameters to be estimated.

4.2.1. Mesh parameterization

In a finite element approach, a mesh $\mathcal{T}_h$ is created in order to approximate the functional space in which the solutions of I_s, I_c, I_s^* and I_c^* are searched. In this study, these functions are searched within the space $V_h(\mathcal{T}_h, \mathcal{F}) := \{v_h \in C^1(\mathcal{D}),\ v_h|_K \in \mathcal{F}\ \forall K \in \mathcal{T}_h\}$ where $\mathcal{F}$ is a space of functions defined later and K denotes a triangle.

When using finite elements methods, it seems interesting to use the same mesh $\mathcal{T}_h$ to construct the approximate functional space for the unknown functions κ and σ, say the space $\mathcal{G}$ (the next subsection deals with such choice for $\mathcal{G}$ with for instance the piecewise constant approximation $\mathcal{G} := \mathcal{P}_0$, the continuous and linear per element approximation $\mathcal{G} := \mathcal{P}_1$ or other functional spaces).

However, as in cases presented in [41], the choice of $\kappa,\ \sigma \in V_h(\mathcal{T}_h, \mathcal{G})$ is likely to give poor results in the reconstruction, because too many parameters (values of $\kappa,\ \sigma$ on the nodes of $\mathcal{T}_h$ or on internal elements of $\mathcal{T}_h$) are dependent. This is due to over-parameterization with respect to the lack of measurement information.

In [43] some elements of proof are given for the number of retrievable parameters. In practice, the singular value decomposition can be performed on the jacobian matrix and the number of retrievable parameters is determined by the rank of the matrix where the noise of the measure gives the limit. Also, [41, 44]

presents the Gauss–Newton matrix condition number fall when using an appropriate re-parameterization of the determination of a heat flux on a one-dimensional boundary.

The parameterization for the parameters can be chosen on a different mesh than $\mathcal{T}_h$, say $\mathcal{M}_h$, yielding to retrieve $\tilde{\kappa},\ \tilde{\sigma} \in V_h(\mathcal{M}_h, \mathcal{G})$ for instance. We thus need to introduce a parameterization map ψ such that $\kappa = \psi(\tilde{\kappa})$, $\sigma = \psi(\tilde{\sigma})$, $(\tilde{\kappa}, \tilde{\sigma}) \in \tilde{C}$ with the inclusion $\psi(\tilde{C}) \subset C$. The (discrete) optimization problem which is actually solved instead of (14) consists thus in finding $(\tilde{\kappa}^*, \tilde{\sigma}^*) = \arg\min j(\psi(\tilde{\theta}))$ over $\tilde{C}$.

Sometimes, the space $V_h(\mathcal{M}_h, \mathcal{G})$ for κ and σ is chosen such that $\mathcal{M}_h \subset \mathcal{T}_h$, that is $\mathcal{T}_h$ is somehow a refinement of $\mathcal{M}_h$ [43]. However, such embedding is not compulsory, the mesh $\mathcal{M}_h$ can be chosen independently of $\mathcal{T}_h$ provided that the number of parameters to retrieve remains acceptable and also provided that the routine $\psi(\,\cdot\,)$ is well implemented within the finite element setting.

In any case, such parameterization, i.e. choosing $\dim V_h(\mathcal{M}_h, \mathcal{G}) < \dim V_h(\mathcal{T}_h, \mathcal{F})$ regularizes the inverse problem by reducing the number of parameters to be estimated while keeping an accurate forward and adjoint models solutions on finer mesh. On one hands this choice reduces the number of unknowns thus stabilizing the optimization. On the other hand, this parameterization also imposes (implicitly) regularity constraints which yields to a stabilizing effect on the inverse problem [43]. Note that sometimes, this methodology is also called the dual mesh approach [45].

4.2.2. Functional space

The other aspect of regularizing through parameterization consists in choosing the appropriate function space $\mathcal{G}$ for searching $\kappa,\ \sigma \in V_h(\mathcal{M}_h, \mathcal{G})$. When searching space-dependent physical properties in a finite elements setting, the most usual space for $\mathcal{G}$ is the piecewise constant approximation $\mathcal{P}_0 = \{v \in L_2(\mathcal{D}) | \forall K \in \mathcal{T}_h, \exists \alpha_K \in \mathbb{R} : v_k = \alpha_K\}$, though this functional space yields to two to three times the number of degree of freedom than with the first-order piecewise continuous finite element $\mathcal{P}_1 = \{v \in H^1(\mathcal{D}) | \forall K \in \mathcal{T}_h, v_k(x, y) = a + bx + cy\}$. Moreover, the space $\mathcal{P}_0 \subset L_2(\mathcal{D})$ does not impose regularity as with continuous space defined within $H^1(\mathcal{D})$. Other functional spaces are used for comparison in the numerical dedicated section. The key point is to choose, for a given mesh $\mathcal{M}_h$, a functional space that yields to, on one hand a not so large number of degrees of freedom and, on the other hand continuity, thus regularity for the solution.

4.3. Noise filtering with cost function gradient in Sobolev space

The source term acting on the adjoint problem is the discrepancy between the measurement and the prediction, i.e. $(P - M)\, P'(I_s)$ applied on $\partial\mathcal{D}_d$. The measurements M are noised in real cases. These noises propagate to the adjoint states I_s^* and I_c^* that results in fluctuations in the cost function gradient $\nabla j(\theta)$ and then to the parameters. When represented in space x, *i.e.* defining $h(x) := \nabla j(\theta)$, such gradient function image presents high frequency non smooth fluctuations whose cause is the measurement noise. Such images of non-smooth cost function gradients are presented in the numerical results dedicated section.

One way to avoid such noisiness, is to filter the image $h(x)$. Among the numerous methods [46], the one based on the Sobolev gradient used for image segmentation [47] is interesting due to its inherent relationships with preconditionning the gradient for optimization [48, 49, 50, 51, 52, 53] and for smoothing [54]. Such gradient has been used recently in heat transfer for estimating time and/or space conditions where the classical cost gradient is null. Then, an alternative Sobolev gradient combined with a Tikhonov-type penalization on the space and time derivatives of the state was introduced [54]. While several inner product definitions have been tested for non linear optimization in [55, 56], we are here only interested in smoothing the cost gradient to filter the noise. To do so, the gradient of the cost function is extracted from the Sobolev space equipped with the inner product [56]:

$$\langle u, v \rangle_{H^{1(\ell)}(\mathcal{D})} := \frac{1}{1+\ell^2} \int_{\mathcal{D}} \left(\bar{u}v + \ell^2 \nabla \bar{u} \cdot \nabla v \right) \, \mathrm{d}x \tag{25}$$

where $\ell \in \mathbb{R}$ is a tuning parameter. The identification $(\nabla^S j, \theta')_{H^{1(\ell)}(\mathcal{D})} = j'(\theta; \theta')$ gives, after integration by parts:

$$\frac{1}{1+\ell^2} \left(1 - \ell^2 \Delta\right) \nabla^S_\kappa j(\theta) = \overline{I_c} I_c^* + \overline{I_s} I_s^* + \epsilon(\kappa - \kappa_0) \tag{26}$$

$$\begin{aligned} &\tfrac{1}{1+\ell^2} \left(1 - \ell^2 \Delta\right) \nabla^S \sigma j(\theta) = \overline{I_c} I_c^* + \overline{I_s} I_s^* + \epsilon(\sigma - \sigma_0) \\ &-\frac{\overline{I_s^*}}{4\pi} \int_{4\pi} \left[I_s(r, \vec{\Omega}', \omega) + I_c(r, \omega) \delta(\vec{\Omega}' - \vec{\Omega}_c) \right] \Phi(\vec{\Omega}', \vec{\Omega}) \, \mathrm{d}\Omega' \end{aligned} \tag{27}$$

where we used the fact that the direction $\theta' = 0$ on $\partial\mathcal{D}$ in the integration by parts.

In practice, the $L_2(\mathcal{D})$ gradient of the cost function is extracted from (A.18)-(A.19), then one computes the Helmoltz problem to extract the so-called Sobolev gradient, with α being either κ or σ:

$$\frac{1}{1+\ell^2} \left(1 - \ell^2 \Delta\right) \nabla^S_\alpha j(\theta) = \nabla_\alpha j(\theta) \tag{28}$$

Note that according to [56], the inverse Helmoltz operator used to extract the Sobolev gradient is, when regarded in the Fourier space, equivalent to a low-pass filter, where the cut-off is related to the inverse of the length scale ℓ. So, the solution of the Helmoltz problems (A.18)-(A.19) de-emphasizes the components with the smallest characteristic length-scales.

5. Tests and results

5.1. Test description

A 2 cm $\times$ 2 cm domain which contains two inclusions is chosen to test an optical tomography reconstruction with the above forward model based on the Discontinuous Galerkin formulation. The values of the optical properties of the medium are given in Table 1 and their simulated distributions are shown in Figure 1. It is assumed that the medium is forward-scattering where the phase function is given by the Henyey-Greenstein function with an anisotropic factor g= 0.9 . Four collimated sources with zero-phased, modulated

at the frequency of 600 MHz are placed at mid-centres of each side of the square and the measurements are done with eight detectors of 0.8 cm of extension regularly placed around the boundaries beyond the source positions.

The synthetic data are generated on a regular fine triangular mesh of 8142 elements (element size of $h = 1/60$) and 24 angular discrete ordinates (S_6 quadrature) by using the above forward model. The original parameters are also generated with the same mesh. In order to not commit the inverse-crime, the state space in the inverse problem is based on a coarser mesh of 3786 elements coupled to 24 discrete directions while the parameters are estimated on different coarser meshes.The generated complex-valued intensities are used as the reference data for the inversion with noise where a Gaussian distribution is used such that the level of the noise in dB unit is given by

$$SNR = 10 \log_{10} \left(\frac{M_d}{\sigma_{M_d}} \right) \tag{29}$$

where σ_{M_d} is the standard deviation.

The inversion computations are performed with the limited memory quasi-Newton method of Brodyen-Fletcher-Goldfarb-Shanno [40, 57]. In order to gauge accuracy of the reconstruction, we introduce the errors e_1, e_2 and e_3 defined by:

$$e_1 = \frac{1}{N_c} \sum_{i=1}^{N_c} \left(\frac{\theta_i^r - \theta_i^o}{\theta_i^o} \right)^2 \tag{30}$$

$$e_2 = \left(\frac{\int_{\mathcal{D}} \left(\theta_i^r - \theta_i^o \right)^2 dx}{\int_{\mathcal{D}} \left(\theta_i^o \right)^2 dx} \right)^{1/2} \tag{31}$$

$$e_3 = \left(\int_{\mathcal{D}} \left(\nabla \theta_i^r - \nabla \theta_i^o \right)^2 dx \right)^{1/2} \tag{32}$$

where N_c is the number of degrees of freedom related to the parameters finite element space, and superscripts r and o refer to the reconstructed and original images, respectively. e_1 represents the mean quadratic errors per degree of freedom, e_2 is the relative error of the reconstruction with respect to the original distribution, e_3 is the relative error of the reconstruction with respect to the original distribution using the H^1 norm to show the regularity of the solution.

Two criteria are used to break the L-BFGS iteration loop. One is based on the stabilization of the cost function, the other being the limitation of the number of BFGS iterations:

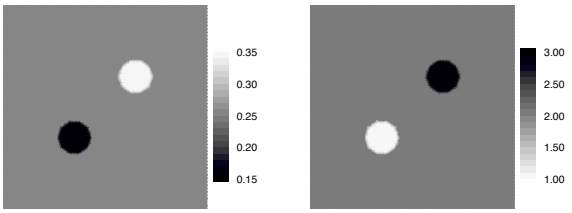
$$\frac{j_k - j_{k-1}}{j_{k-1}} \leq \varepsilon \tag{33}$$

$$k \leq k_{\max} \tag{34}$$

where the user-defined value ε has been fixed to 10^{-3}, and $k_{\max}$ has been fixed to 100 for the computational time purpose as using the full RTE as a forward model in optical tomography with finite element method is

Table 1: Optical properties of the test medium

	Background	Bottom inclusion	Top inclusion
κ	0.25 cm^{-1}	0.35 cm^{-1}	0.15 cm^{-1}
σ_s	20 cm^{-1}	30 cm^{-1}	10 cm^{-1}



(a) (b)

Figure 1: Original distribution of the optical properties:(a) the absorption coefficient, (b) the reduced scattering coefficient.

time consuming. Note that the criterion Eq. (33) is used in order to handle the large gradient values around the collimated sources which contribute to increase its norm [23].

In the following, a series of test is performed in order to compare the different strategies detailed in Section 4. All comparisons are performed with respect to the particular test where κ and σ are in $V_h(\mathcal{M}_h, \mathcal{G})$, the mesh $\mathcal{M}_h$ being composed of 3786 triangles ($h/40$), and $\mathcal{G}$ being the piecewise continuous finite elements $\mathcal{P}_1$. The subsection 5.2.1 compares the identifications results while using different meshes.

The subsection 5.2.2 compares the identifications results while using different functional spaces. The subsection 5.3 deals with the use of the Sobolev inner product definition for extracting the cost function gradient, with different values for the parameter ℓ^2. Then, the subsection 5.4 deals with the use of the classical Tikhonov regularization strategy with different weight for ϵ.

5.2. Finite element parameterization

5.2.1. Parameterization through mesh coarsening

This subsection deals with the use of different meshes $\mathcal{M}_h$ for the approximation of the continuous optical properties. From a large number of experiments, only the most significant results are presented here. The mesh sizes are related to values $h/40$, $h/20$ and $h/10$. Piecewise continuous finite elements are used. There

is no Tikhonov regularization, and the cost function gradient is extracted using the classical $L_2(\mathcal{D})$ inner product defined by (A.10).

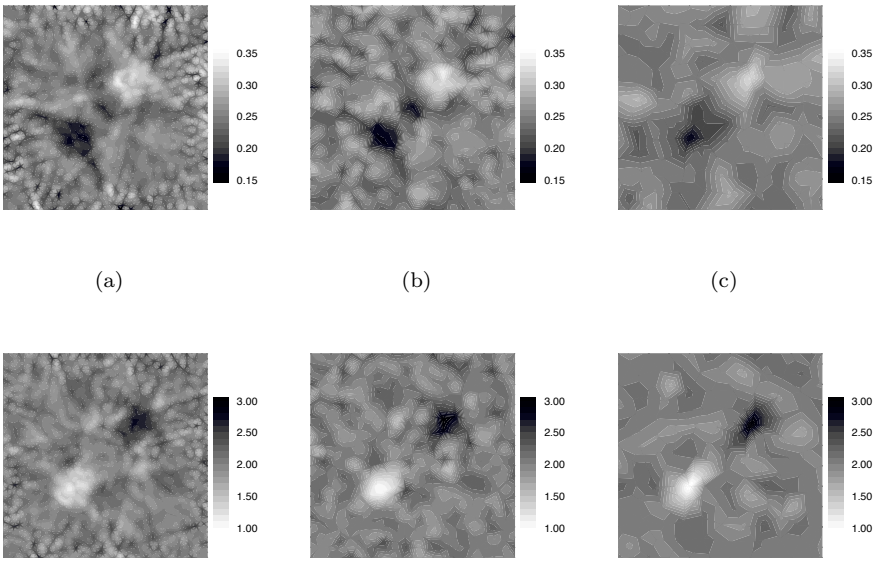
Table 2 gives quantitatively the quality of the identification results, for the four selected meshes. For each, Table 2 gives the number of degrees of freedom, the errors e_1, e_2 and e_3 for each component κ and σ separately, and the decrease of the ratio of the final cost function with respect to the initial ones. Figure 2 represents the 2-dimensional reconstructions and Figure 3 represents the corresponding cuts along the line passing by the bottom-left corner to the top-right corner.

From Table 2, it is seen that the greater is the number of degrees of freedom, i) the smaller is the final cost function value, ii) the smaller is the mean quadratic error $e_{1,\alpha}$, iii) the greater is the error $e_{3,\alpha}$ related to the regularity, iv) while the integral relative error $e_{2,\alpha}$ remains more or less stable. Figures 2 and 3 confirm these analyses: a coarse mesh is needed to obtain a regular solution while a fine fine is needed to get as close as the objective. Taking into account of the whole results (errors and curves) related to the finite elements meshes for the optical properties, intermediate meshes with element sizes around $h/20$ are the ones that fulfill the compromise between regularity and bias with respect to the original distributions. On the opposite, the coarser mesh does not fulfill the small quadratic and integral errors, and the finer mesh does not fulfill regularity. Note that the whole number of unknowns are related to the mesh size and to the used finite elements space order. This is the purpose of section 5.2.2.

5.2.2. Parameterization with different functional spaces

In this section, an analysis with a finite element parameterization of the space of the unknowns is done through a comparative test of reconstruction with different finite element spaces for the optical properties where the intensity fields are supposed to be in the space of linear discontinuous elements ($\mathcal{P}_{1\mathrm{dc}}$). All the reconstructions are carried out with the computed noised data. The reconstructions are done by taking the parameters respectively in the space of piecewise constant ($\mathcal{P}_0$), piecewise linear discontinuous ($\mathcal{P}_{1\mathrm{dc}}$), piecewise linear continuous ($\mathcal{P}_1$) and piecewise quadratic ($\mathcal{P}_2$) elements. The recovered distribution are reported in Figure 4 for both properties. A cut alone the line passing by the bottom-left corner to the top right corner of the geometry is given in Figure 5.

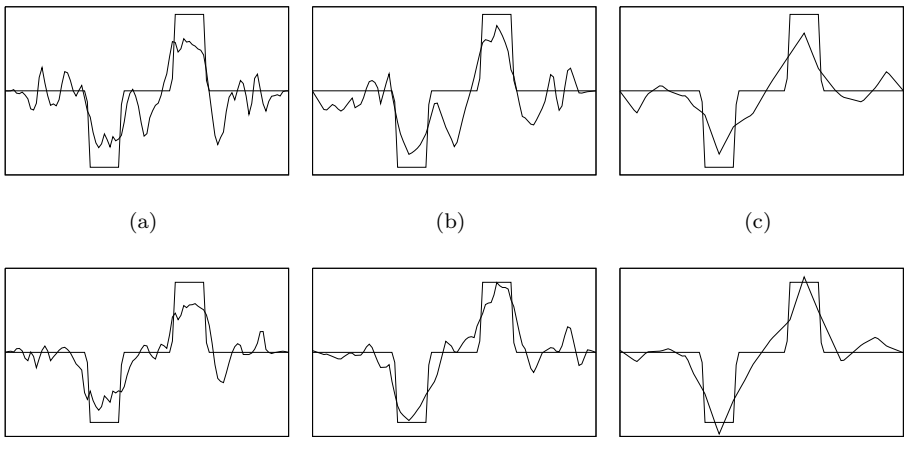
From the results, it is seen that the use of continuous finite elements spaces gives better results than the discontinuous one. The results show that continuous finite elements approximations of the optical properties introduce some implicit regularization on the inversion by smoothing the results. For each type of finite element space (continuous or discontinuous), it is seen that increasing the number of degrees of freedom do not improve the quality of the reconstruction. The smallest errors defining the quality of the reconstruction is given by the $\mathcal{P}_1$ finite element space whose number of degrees is the lowest (see ϵ_2 in Table 3). The errors given by ϵ_1 take into account the number of degrees of freedom and show that increasing the number of degrees of freedom lower the errors ϵ_1 as expected. However, taking into account the total number of degrees



(a) (b) (c)

(d) (e) (f)

Figure 2: Mesh parameterization results. 2-D reconstruction. Top row the absorption coefficients, bottom row, the reduced scattering coefficients. (a) and (d) results with $h = 1/40$, (b) and (e) results with $h = 1/20$,(c) and (f) results with $h = 1/10$ where h is the mean size of the mesh element where the optical properties are searched for.



(a) (b) (c)

(d) (e) (f)

Figure 3: Mesh parameterization results. 1-D reconstruction on the line passing by the bottom-left and the top-right corners. Top row the absorption coefficients, bottom row, the reduced scattering coefficients. (a) and (d) results with $h = 1/40$, (b) and (e) results with $h = 1/20$,(c) and (f) results with $h = 1/10$, h is the mean size of the mesh element where the optical properties are searched for.

Table 2: Comparative accuracy of the reconstruction for different meshes. e_i, $i = 1, 2, 3$ are the errors defined by (30), (31) and (32). , $ndof$ is the number of degree of freedom of the finite element space, j_0 and j_f are respectively the initial and the final cost function values.

		$h = 1/40$	$h = 1/20$	$h = 1/10$
$ndof$	–	3952	1006	310
e_1	$e_{1,\kappa}$	0.0026	0.0048	0.0114
	$e_{1,\sigma}$	0.0027	0.0050	0.0157
e_2	$e_{2,\kappa}$	0.0893	0.0805	0.0806
	$e_{2,\sigma}$	0.0889	0.0764	0.0919
e_3	$e_{3,\kappa}$	1.6328	1.0195	0.8787
	$e_{3,\sigma}$	13.036	9.2436	8.4539
j_f/j_0	–	0.0255	0.0640	0.0982

of freedom, the piecewise linear elements ($\mathcal{P}_1$) solution remains the more accurate.

From the inverse analysis point of view, high order parameterization may increase the ill-posed nature of the inverse problem such as the existence of correlations, the number of possible solutions and the sensitivity to round errors and noise. Then, increasing the number of degrees of freedom of the space of the unknowns, one may end-up with less accurate solutions due to the increase of the ill-posed behavior of the inverse problem. However reducing the number of unknowns is a way of reducing the number of possible solutions *i.e* a way of regularizing the inversion or in other word, reducing the ill-posed behavior of the problem.

Physically, with continuous parameterization, information is well conducted among the whole computational nodes which improves the reconstruction as seen in the results. Indeed, when discontinuous fields are used for the space of the unknowns in the inverse approach for the parameters, there is no connection between them as in direct simulation where numerical fluxes are used to ensure connection between the whole computation domain. Then, continuous parameterization has two advantage in reducing the number of unknown and keeping connection for a better transferring of information for a good estimation.

Concluding with the used finite elements spaces for the parameters, the continuous parameterization in the space $\mathcal{P}_1$ is the one that fulfill both properties of f i) giving the lower number of degrees of freedom, ii)

while keeping continuity to transfer of information.

Table 3: Comparative accuracy of the reconstruction for different finite element spaces. e_i, $i = 1, 2, 3$ are the errors defined by (30), (31) and (32). $ndof$ is the number of degree of freedom of the finite element space, j_0 and j_f are respectively the initial and the final cost function values.

		$\mathcal{P}_0$	$\mathcal{P}_{1\mathrm{dc}}$	$\mathcal{P}_1$	$\mathcal{P}_2$
$ndof$	–	7580	22740	3952	15482
ϵ_1	$\epsilon_{1,\kappa}$	0.0015	0.0008	0.0018	0.0009
	$\epsilon_{1,\sigma}$	0.0018	0.0010	0.0021	0.0010
ϵ_2	$\epsilon_{2,\kappa}$	0.0932	0.0895	0.0651	0.0720
	$\epsilon_{2,\sigma}$	0.1005	0.0964	0.0715	0.0770
e_3	$e_{3,\kappa}$	1436	4159	641	2275
	$e_{3,\sigma}$	24552	70310	10680	39109
j_f/j_0	–	0.0136	0.0145	0.0465	0.0309

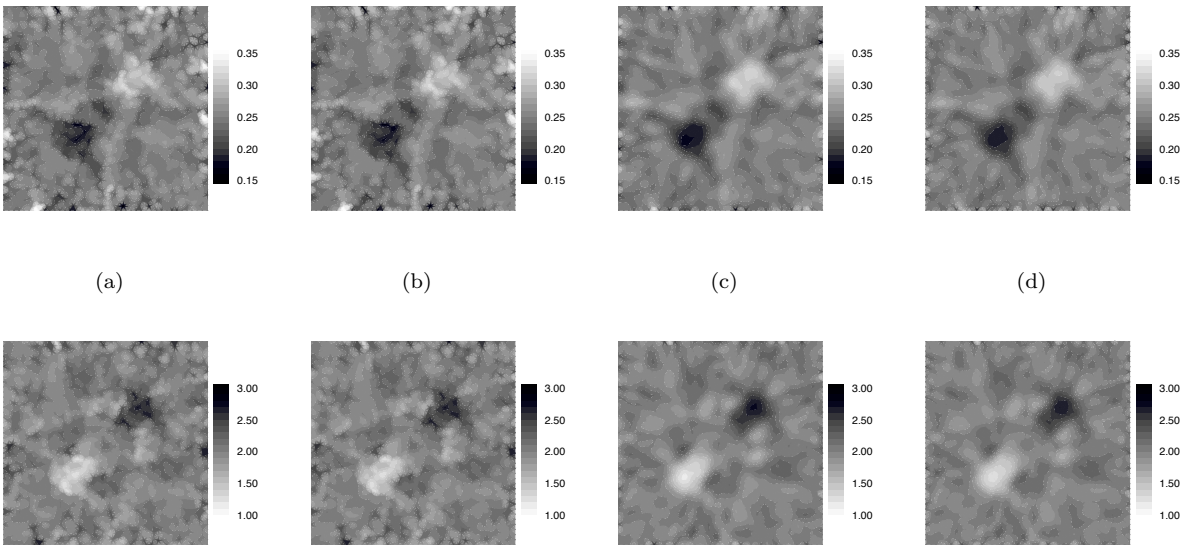


Figure 4: FE space parameterization results. 2-D reconstruction. Top row the absorption coefficients, bottom row, the reduced scattering coefficients. (a) and (e) results with $\mathcal{P}_0$, (b) and (f) results with $\mathcal{P}_{1\mathrm{dc}}$,(c) and (g) results with $\mathcal{P}_1$,(d) and (h) results with $\mathcal{P}_2$.

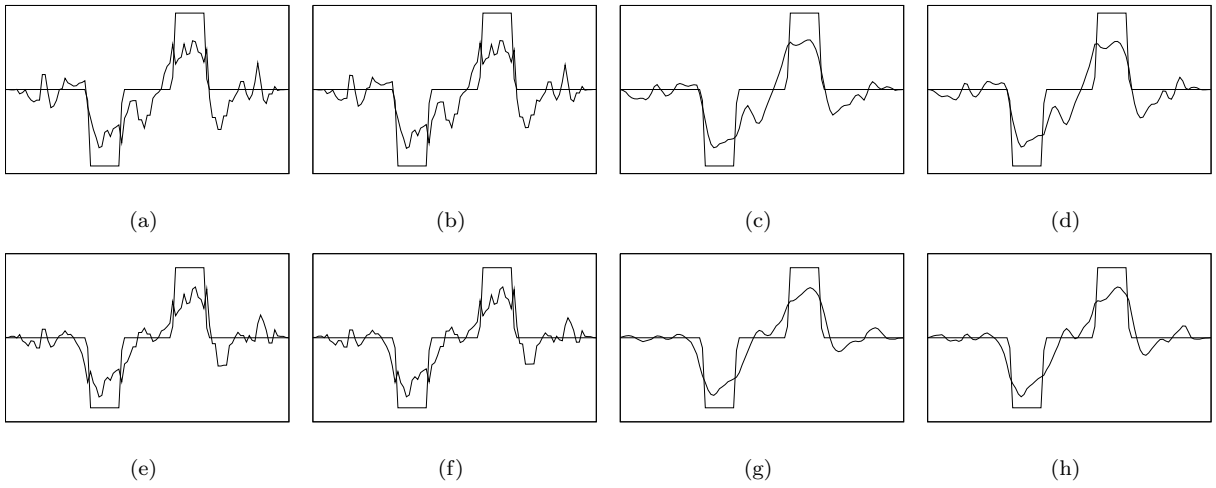


Figure 5: FE space parameterization results. 1-D reconstruction on the line passing by the bottom-left and the top-right corners. Top row the absorption coefficients, bottom row, the reduced scattering coefficients. (a) and (e) results with $\mathcal{P}_0$, (b) and (f) results with $\mathcal{P}_{1\mathrm{dc}}$,(c) and (g) results with $\mathcal{P}_1$,(d) and (h) results with $\mathcal{P}_2$.

5.3. Results with Sobolev cost function gradient

This subsection deals with the use of the Sobolev inner product definition (25) to extract the cost function gradient. From a large number of experiments, only the most significant results are presented here. The size of the mesh is fixed to $h/40$ and first-order piecewise continuous finite elements approximations are used. There is no Tikhonov regularization. In this presentation, the parameter ℓ^2 has been fixed to three values within the range $[10^{-4}, 5 \times 10^{-3}]$.

Table 4 gives quantitatively the quality of the identification results, for the three selected smoothing values ℓ^2. For each, Table 2 gives the number of degrees of freedom, the errors e_1, e_2 and e_3 for each component κ and σ separately, and the decrease of the ratio of the final cost function with respect to the initial ones.

Figures 6 and 7 represent the 2-dimensional reconstructions and the corresponding cuts along the line passing by the bottom-left corner to the top-right corner.

From Table 2, it is seen that as the Sobolev parameter ℓ^2 increases, i) the final cost function value also increases but at the same time, ii) errors $e_{1,\alpha}$ and $e_{2,\alpha}$ slightly decrease. iii) moreover, there is a large decrease of the error $e_{3,\alpha}$. From Figures 2 and 3, it can be seen that the use of the Sobolev inner product to extract the gradient smoothes the cost function gradient and thus the recovered properties. With a two small Sobolev parameter, there is not enough diffusion of the cost function gradient and so the images of Figures 6(a) and 6(d) are like the ones of the previous subsections with similar parameters. On the opposite, with two high values for the Sobolev parameter, there is too much diffusion of the cost function gradient, yielding to a great attenuation of the fluctuations, but the nominal values of the optical properties are not well recovered. From all tested Sobolev parameter values, and for this presented particular test, values of ℓ^2 around $= 10^{-3}$ is a judicious choice for assuring a great decrease of the cost function, small errors in the properties reconstruction while keeping regularity on the reconstruction.

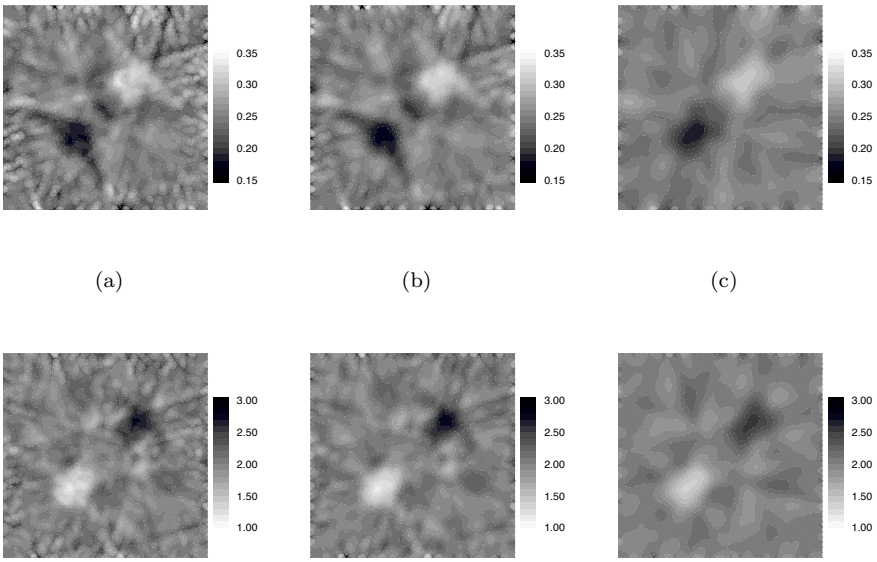


Figure 6: Results with Sobolev gradient: Top row the absorption coefficients, bottom row, the reduced scattering coefficients. (a) and (d) results with $\ell^2 = 0.0001$, (b) and (e) results with $\ell^2 = 0.001$,(c) and (f) results with $\ell^2 = 0.005$.

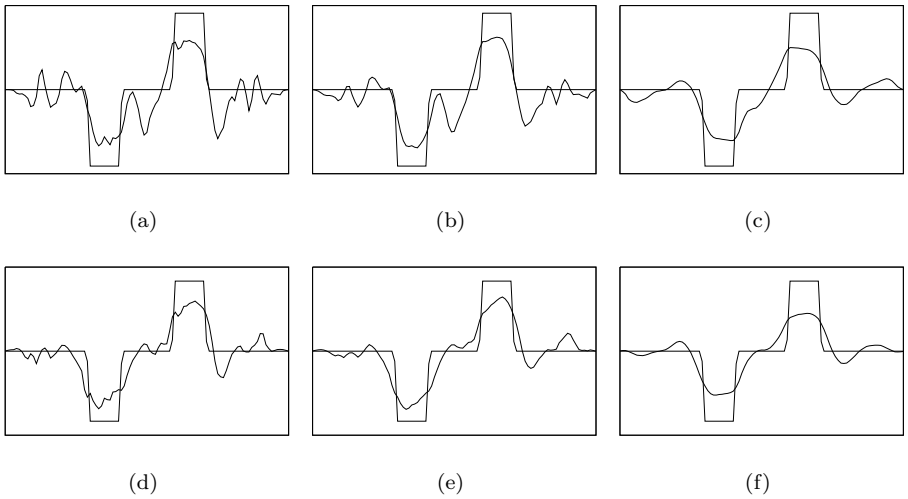


Figure 7: Results with Sobolev gradients. Top row the absorption coefficients, bottom row the reduced scattering coefficients. (a) and (d) with $\ell^2 = 0.0001$, (b) and (e) with $\ell^2 = 0.001$, (c) and (f) with $\ell^2 = 0.005$.

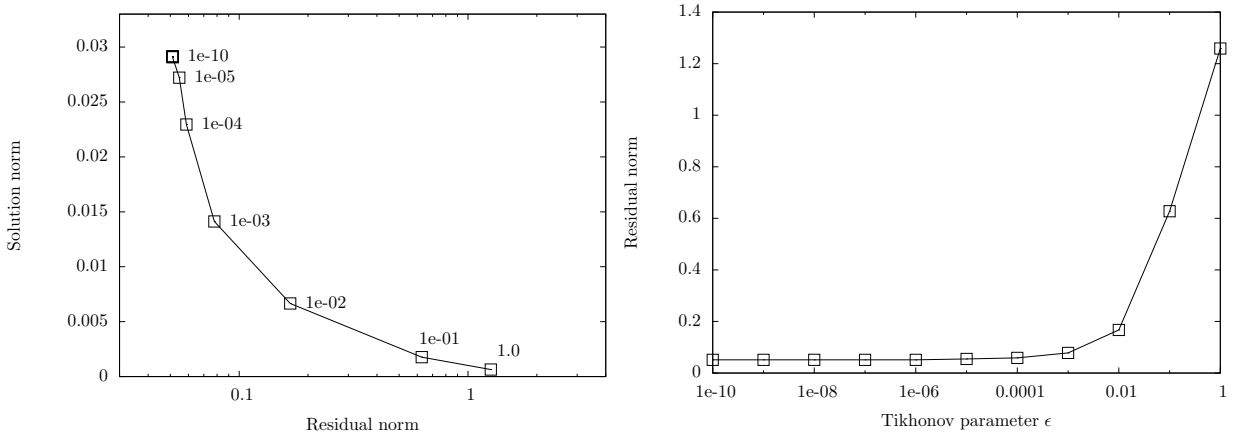
Table 4: Comparative accuracy of the reconstruction for different values for the smoothing parameter ℓ^2. e_i, $i = 1, 2, 3$ are the errors defined by (30), (31) and (32). $ndof$ is the number of degree of freedom of the finite element space, j_0 and j_f are respectively the initial and the final cost function values.

		$\ell^2 = 0.0001$	$\ell^2 = 0.001$	$\ell^2 = 0.005$
$ndof$	–	3952	3952	3952
e_1	$e_{1,\kappa}$	0.0023	0.0021	0.0017
	$e_{1,\sigma}$	0.0025	0.0022	0.0021
e_2	$e_{2,\kappa}$	0.0825	0.0744	0.0621
	$e_{2,\sigma}$	0.0832	0.0745	0.0699
e_3	$e_{3,\kappa}$	1.36255	1.16738	0.877141
	$e_{3,\sigma}$	11.3714	9.77341	8.36711
j_f/j_0	–	0.029	0.0377	0.0751

5.4. Results with Tikhonov Regularization

This subsection deals with the use of the Tikhonov penalty regularization technique. From a large number of experiments, only the most significant results are presented here: a test with a too low value for ϵ, a case with a too high value for ϵ, and a case with an appropriate value for ϵ chosen on the corner of the L-curve. The size of the mesh is fixed to $h/40$ and first-order piecewise continuous finite elements approximations are used. The cost function extraction is performed according to the $L_2(\mathcal{D})$ inner product definition. The value for the weight ϵ varies within the range $[10^{-10}, 10^{-1}]$ and the corresponding L-curve with the evolution of the residual norm are given in Fig. 8. It is seen that the L-curve (see Fig. 8(a)) has a L-shape as expected. However, the corner is not distinctive in the corresponding test due to difficulties induced by the nonlinearities, the different order of magnitude between κ and σ. Let us remark that such cases are reported in the literature for nonlinear inverse problem [58].

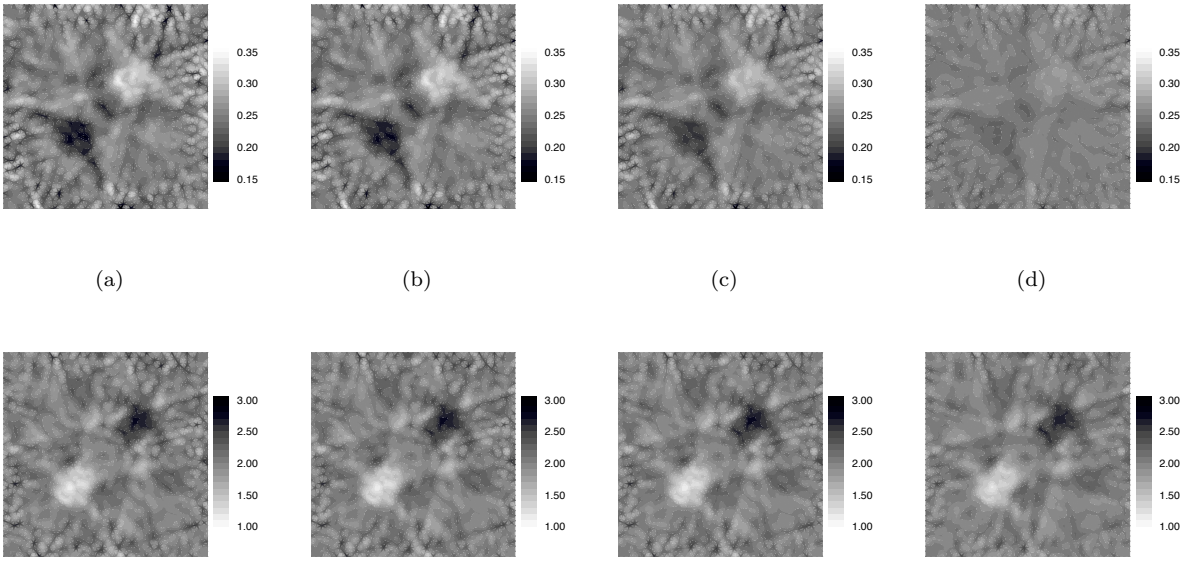
Then selected results are given in Table 5 according to the possible L corner of the L-curve where the corresponding estimated distribution are plotted in Fig 9. It is seen when the regularization parameters ϵ is increasing the errors e_1, e_2 and e_3 are decreasing while j_f/j_0 is increasing. The estimated distributions (Fig. 9) and the cross line plots (Fig. 10) show that good accuracy is obtained for $\epsilon = 10^{-4}$ and $\epsilon = 10^{-5}$.



(a) (b)

Figure 8: Analysis of the Tikhonov parameter ϵ. (a) L-curve for the Tikhonov test with $\alpha \in [10^{-10}, 10^{-1}]$. (b) evolution of the residual with respect to ϵ. The residual and the solution norm are given respectively by $\mathcal{J}(I_s)$ and $\|M(\theta - \theta_0)\|$ where M is a diagonal matrix such that $\|M(\theta - \theta_0)\| = \dfrac{\|\kappa - \kappa_0\|}{\|\kappa_0\|} + \dfrac{\|\sigma - \sigma_0\|}{\|\sigma_0\|}$.

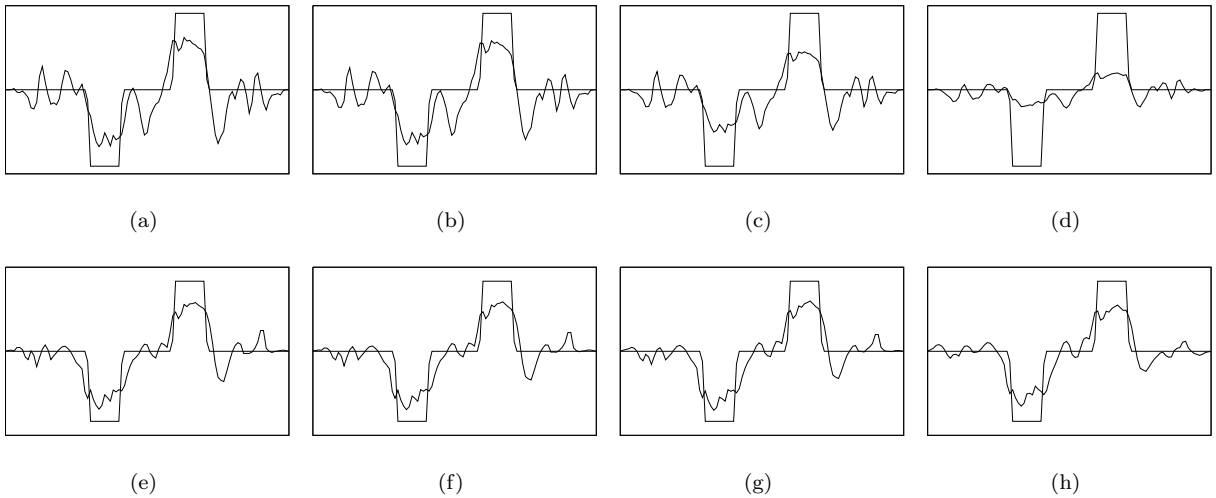
It is seen that the Tikhonov regularization tends to reduce the oscillations in the reconstructions (see Fig. 10) where higher value of the regularization parameter.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 9: Results with Tikhonov regularization: Top row the absorption coefficients, bottom row, the reduced scattering coefficients. (a) and (e) results with $\epsilon = 0$, (b) and (f) results with $\epsilon = 10^{-5}$,(c) and (g) results with $\epsilon = 10^{-4}$, (d) and (h) results with $\epsilon = 10^{-3}$.



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 10: Results with Tikhonov regularization. Top row the absorption coefficients, bottom row the reduced scattering coefficients. (a) and (e) results with $\epsilon = 0$, (b) and (f) results with $\epsilon = 10^{-5}$,(c) and (g) results with $\epsilon = 10^{-4}$, (d) and (h) results with $\epsilon = 10^{-3}$.

Table 5: Comparative accuracy of the reconstruction for different values for the Tikhonov parameter weight ϵ. e_i, $i = 1, 2, 3$ are the errors defined by (30), (31) and (32). $ndof$ is the number of degree of freedom of the finite element space, j_0 and j_f are respectively the initial and the final cost function values.

		$\epsilon = 0$	$\epsilon = 10^{-5}$	$\epsilon = 10^{-4}$	$\epsilon = 10^{-3}$
e_1	$e_{1,\kappa}$	0.0893	0.0859	0.08029	0.07507
	$e_{1,\sigma}$	0.0889	0.0871	0.0854	0.08078
e_2	$e_{2,\kappa}$	0.0026	0.0025	0.00237	0.0022
	$e_{2,\sigma}$	0.0027	0.0026	0.00258	0.0024
e_3	$e_{3,\kappa}$	1.63287	1.5257	1.3714	1.0375
	$e_{3,\sigma}$	13.0364	12.4958	12.0032	10.5521
j_f/j_0	–	0.0255	0.0273	0.0294	0.0389713

6. Comparative analysis of the regularization tools

From a comparative analysis of the results, it is seen that parameterization of the space of the unknown parameters with the mesh size and the order of finite element approximation aim at lowering the number of degree of freedom of the parameters and reduce the ill-posed behavior of the inverse problem as seen in [42] . This approach have shown that continuous approximation in the case of finite element framework is more appropriate as this allows a connection among the whole computational nodes which improves the reconstruction as seen in the results. Indeed, when discontinuous fields are used, there is no way to transfer information between the nodes during the inverse approach contrary to direct approach where numerical fluxes can be used.

The analysis show that the use of the filtering through the "Sobolev gradient" allow to filter noise in the construction by smoothing the results where an optimal filtering parameter is to be determined. Compare to the Tikhonov approach, the filtering by "Sobolev gradient" gives more smoothing and regularity in the solution with a easy setting. However, an additional equations is to be solve during each iteration step of the gradient based algorithm.

The presents results show that the Tikhonov approach reduces the oscillations of the estimated distributions but does not enhance regularity of the solution as the Sobolev gradient does when the data is noisy.

Moreover, the choice of the optimal Tikhonov weight parameter remains problematic since a large number of inverse problems runs has to be performed when searching this optimal parameter particularly in nonlinear cases where the L-curve corner maybe non distinctive [58].

7. Conclusion

Various numerical tools have been developed for improving the quality and the contrast of the estimated distribution of the optical properties in optical tomography applications. The Radiative Transfer Equation in frequency domain and its adjoint version are approximated according to the Discontinuous Galerkin Finite Elements Methods coupled with the Discrete Ordinate Method. The inverse problem involved in optical tomography is ill-posed in the sense that small measurement errors may cause large errors in the optical properties reconstructions, even when using optimizers that do not rely on some matrix inversion (e.g. Non Linear Conjugate Gradient, BFGS, etc.).

This paper showed new strategies to improve the accuracy of the reconstruction through mesh and finite elements parameterization of the optical properties to be recovered, the use of gradient filtering through an alternative inner product instead of the classical L_2 inner product within the adjoint method. The paper gives of a more appropriate inner product definition such as the Sobolev one in order to extract a smoother cost function gradient yielding to a better reconstruction of the optical properties with higher regularity.

The numerical experiments also showed that the classical use of the same finite elements mesh for the computation of both, the state (forward and adjoint) variables and the solution of the inverse problem (here the optical properties) yielded to less accurate reconstructions. Indeed, when doing so, we are faced to over-parameterization with respect to the lack of measurement information whose effect is the increase of the ill-posed behavior of the inverse problem. Then, a used of corse mesh for the optical parameter with finer mesh for the solution of the forward model, also called dual mesh approach is more appropriate.

The study showed that the number of parameters must be lowered using coarser meshes for the description of the optical properties coupled with finite elements spaces of low degrees of freedom. In addition, continuous finite element space gives better results as they have a low number of degrees of freedom and information is well conducted among the whole computational nodes which improves the reconstruction.

These regularization strategies have been confronted to the use of the Tikhonov regularization. The comparisons showed that the Tikhonov regularization actually de-emphasizes the fluctuations on the results, but the regularity and the contrast are lower than when using an appropriate re-parameterization on one hand, and the Sobolev cost function gradient on the other hand.

Future studies will show an efficient way of choosing this parameters according to mesh element size for a better implementation of the filtering parameters with an optimal parameterization approach. Also, the optimization algorithm will takes advantages of the Sobolev gradient in order to improve the minimization

algorithm with a trust-region method approach [59].

Acknowledgement

This work was supported by the National Sciences and Engineering Research Council of Canada (NSERC). The authors would like to thank Pr. Charette and J. Boulanger for their valuable introduction to optical tomography. They would also like to thank Pr. Jarny for interesting scientific discussions and for introducing to them the so-called "Sobolev gradient".

Appendix A. Gradient computation with the adjoint method

Below we show the process for the deduction of the cost function gradient where the forward model equations are taken as an implicit constraint to the minimization problem of $j(\theta)$. For simplicity in the following mathematical formulations, let us rewrite the forward model equations (Eq. (4) and Eq.(6)) as

$$(\boldsymbol{a} + \boldsymbol{b}(\sigma, \kappa))\, I_c = 0 \tag{A.1}$$

$$(\boldsymbol{a} + \boldsymbol{b}(\sigma, \kappa))\, I_s = \boldsymbol{c}(\sigma)\boldsymbol{f}(I_c, Is) \tag{A.2}$$

where

$$\begin{aligned}
\boldsymbol{f} &:= \frac{1}{4\pi}\int_{4\pi}\left[I_s(r,\vec{\Omega}',\omega) + I_c(r,\omega)\delta(\vec{\Omega}' - \vec{\Omega}_c)\right]\Phi(\vec{\Omega}',\vec{\Omega})\,\mathrm{d}\Omega', \\
\boldsymbol{a} &:= \left(\vec{\Omega}\cdot\nabla + \frac{i\omega}{c}\right), \\
\boldsymbol{b} &:= (\kappa + \sigma), \\
\boldsymbol{c} &:= \sigma.
\end{aligned} \tag{A.3}$$

By definition, the directional differential of the cost function j' in the neighborhood of θ, in the feasible direction θ' writes [60, 61, 62]

$$j'(\theta;\theta') = \lim_{\varepsilon\to 0}\frac{j(\theta+\varepsilon\theta') - j(\theta)}{\varepsilon} \tag{A.4}$$

and the solution of the optimization problem (14) is obtained iteratively until theoretically $j'(\theta;\theta') = 0$, and practically when $j'(\theta;\theta') \leq \epsilon\ \forall\theta'$. Other criteria may be used, either based on the cost function value, or on the cost function gradient norms. The differential of the cost function j in the direction θ' is given by:

$$j'(\theta;\theta') = \cancel{\Re}\left[\overline{(P(I_s) - M)\,P'(I_s)}, I_s'(\theta;\theta')\right]_{Re} + \{\epsilon(\theta-\theta_0),\theta'\}_{\mathcal{U}} \tag{A.5}$$

where $P'(I_s)$ is the partial derivative of the prediction function with respect to I_s, and I_s' is the directional derivative of the state I_s, at the point θ toward the direction θ', and $[z]_{Re} = \Re(z)\ \forall z \in \mathbb{C}$ is the real part of the complex number z. For conciseness, this cost function differential is also denoted afterwards as:

$$j'(\theta;\theta') = [(P-M)P', I_s']_{L_2(\partial D_d)} + \{\epsilon(\theta-\theta_0),\theta'\}_{\mathcal{U}} \tag{A.6}$$

with the corresponding inner product definition:

$$[u,v]_{L_2(\partial D_d)} := \int_{\mathcal{D}} \bar{u}v\,\mathrm{d}x \tag{A.7}$$

The directional derivatives I_c' and I_s' of respectively the collimated component I_c and the scattered

component I_s at the point θ toward the direction θ' are given by the linear tangent problem [63]:

$$\begin{cases} \left(\boldsymbol{a}+\boldsymbol{b}(\sigma,\kappa)\right) I_c' + \boldsymbol{b}'\theta' I_c = 0 & \forall x \in \mathcal{D} \\ I_c' = 0 & \forall x \in \partial\mathcal{D}^- \\ \left(\boldsymbol{a}+\boldsymbol{b}(\sigma,\kappa)\right) I_s' + \boldsymbol{b}'\theta' I_s - \boldsymbol{c}'\theta' \boldsymbol{f} - \boldsymbol{c}\left(\boldsymbol{f}'_{I_c} I_c' + \boldsymbol{f}'_{I_s} I_s'\right) = 0 & \forall x \in \mathcal{D} \\ I_s' = 0 & \forall x \in \partial\mathcal{D}^- \end{cases} \tag{A.8}$$

where $\boldsymbol{b}'\theta'$ (resp. $\boldsymbol{c}'\theta'$) is the derivative of $\boldsymbol{b}$ (resp. $\boldsymbol{c}$) in the direction θ', and $\boldsymbol{f}'_{I_s}$ (resp. $\boldsymbol{f}'_{I_c}$) is the partial derivative of $\boldsymbol{f}$ with respect to I_s (resp. I_c), and $\partial\mathcal{D}^-$ denotes the inflow boundary defined according to the direction of light propagation $\vec{\Omega}$, i.e. $\partial\mathcal{D}^- = \left\{x \in \partial\mathcal{D},\ \vec{\Omega}\cdot\vec{n} < 0\right\}$.

The Lagrange function is formally defined as $\mathbf{L}(I_c, I_s, I_c^*, I_s^*, \theta)$:

$$\begin{aligned} \mathbf{L}(\,\cdot\,) := \quad & \mathcal{J}(I_s) + \mathcal{J}^+(\theta) \\ & + \langle \left(\boldsymbol{a}+\boldsymbol{b}(\sigma,\kappa)\right) I_c, I_c^* \rangle_{L_2(\mathcal{D})} \\ & + \langle \left(\boldsymbol{a}+\boldsymbol{b}(\sigma,\kappa)\right) I_s - \boldsymbol{c}\boldsymbol{f}(I_c, I_s), I_s^* \rangle_{L_2(\mathcal{D})} \\ & + \left\langle I_c - q_0 \mathbb{1}_{[r\in\partial\mathcal{D}_0]}, I_c^* \right\rangle_{L_2(\partial\mathcal{D}^-)} \\ & + \langle I_s, I_s^* \rangle_{L_2(\partial\mathcal{D}^-)} \end{aligned} \tag{A.9}$$

where I_c^* (resp. I_s^*) are the adjoint (co-states) of I_c (resp. I_s), and we have the inner product definition:

$$\langle u, v \rangle_{L_2(\mathcal{D})} := \int_{\mathcal{D}} \bar{u} v \,\mathrm{d}x \quad \text{and} \quad \langle u, v \rangle_{L_2(\partial\mathcal{D}^-)} := \int_{\partial\mathcal{D}^-} \bar{u} v \,\mathrm{d}s \tag{A.10}$$

The differentiated Lagrange function with respect to θ in the direction θ', i.e. $\mathbf{L}'_\theta(I_c, I_s, I_c^*, I_s^*, \theta; \theta')$ is:

$$\begin{aligned} \mathbf{L}'_\theta(\,\cdot\,;\theta') = \quad & [\mathcal{J}'(I_s), I_s']_{L_2(\partial\mathcal{D}_d)} + \{\epsilon(\theta-\theta_0), \theta'\}_{\mathcal{U}} \\ & + \langle \left(\boldsymbol{a}+\boldsymbol{b}(\sigma,\kappa)\right) I_c', I_c^* \rangle_{L_2(\mathcal{D})} + \left\langle \boldsymbol{b}'\theta' I_c, I_c^* \right\rangle_{L_2(\mathcal{D})} \\ & + \left\langle \left(\boldsymbol{a}+\boldsymbol{b}(\sigma,\kappa) - \boldsymbol{c}\boldsymbol{f}'_{I_s}\right) I_s', I_s^* \right\rangle_{L_2(\mathcal{D})} - \left\langle \boldsymbol{c}\boldsymbol{f}'_{I_c} I_c', I_s^* \right\rangle_{L_2(\mathcal{D})} \\ & + \left\langle \boldsymbol{b}'\theta' I_s - \boldsymbol{c}'\theta'\boldsymbol{f}, I_s^* \right\rangle_{L_2(\mathcal{D})} + \langle I_c', I_c^* \rangle_{\partial\mathcal{D}^-} + \langle I_s', I_s^* \rangle_{L_2(\partial\mathcal{D}^-)} \end{aligned} \tag{A.11}$$

Transposing operators within inner products and re-ordering, (A.11) is rewritten to:

$$\begin{aligned}\mathbf{L}'_\theta(\,\cdot\,;\theta') = \quad & [\mathcal{J}'(I_s), I'_s]_{L_2(\partial\mathcal{D}_d)} + \{\epsilon(\theta-\theta_0),\theta'\}_{\mathcal{U}} \\ & + \left\langle \left(\boldsymbol{a}+\boldsymbol{b}(\sigma,\kappa)-\boldsymbol{c}\boldsymbol{f}'_{I_s}\right)^* I_s^*, I'_s \right\rangle_{L_2(\mathcal{D})} \\ & + \langle I_s^*, I'_s\rangle_{L_2(\partial\mathcal{D}^-)} + \left\langle \left(\boldsymbol{a}+\boldsymbol{b}(\sigma,\kappa)\right)^* I_c^*, I'_c \right\rangle_{L_2(\mathcal{D})} \\ & - \left\langle \left(\boldsymbol{c}\boldsymbol{f}'_{I_c}\right)^* I_s^*, I'_c \right\rangle_{L_2(\mathcal{D})} + \langle I_c^*, I'_c\rangle_{L_2(\partial\mathcal{D}^-)} \\ & + \left\langle \boldsymbol{b}'\theta' I_c, I_c^* \right\rangle_{L_2(\mathcal{D})} + \left\langle \boldsymbol{b}'\theta' I_s - \boldsymbol{c}'\theta'\boldsymbol{f}, I_s^* \right\rangle_{L_2(\mathcal{D})}\end{aligned} \tag{A.12}$$

where operators of the type u^* are the adjoint of u. To compute the gradient of the cost function without having to compute the sensitivities as they are computationally expensive, the adjoint variables (I_c^* and I_s^*) are chosen particularly such that:

$$\begin{cases} [\mathcal{J}'(I_s), I'_s]_{L_2(\partial\mathcal{D}_d)} + \left\langle \left(\boldsymbol{a}+\boldsymbol{b}(\sigma,\kappa)-\boldsymbol{c}\boldsymbol{f}'_{I_s}\right)^* I_s^*, I'_s \right\rangle_{L_2(\mathcal{D})} + \langle I_s^*, I'_s\rangle_{L_2(\partial\mathcal{D}^-)} = 0 \\ \left\langle \left(\boldsymbol{a}+\boldsymbol{b}(\sigma,\kappa)\right)^* I_c^*, I'_c \right\rangle_{L_2(\mathcal{D})} - \left\langle \left(\boldsymbol{c}\boldsymbol{f}'_{I_c}\right)^* I_s^*, I'_c \right\rangle_{L_2(\mathcal{D})} + \langle I_c^*, I'_c\rangle_{L_2(\partial\mathcal{D}^-)} = 0 \end{cases} \tag{A.13}$$

so that one obtains, by identification:

$$j'(\theta;\theta') = (\nabla j(\theta),\theta') = \{\epsilon(\theta-\theta_0),\theta'\}_{\mathcal{U}} + \left\langle \boldsymbol{b}' I_c I_c^*, \theta' \right\rangle_{L_2(\mathcal{D})} + \left\langle \boldsymbol{b}' I_s I_s^* - \boldsymbol{c}'\boldsymbol{f} I_s^*, \theta' \right\rangle_{L_2(\mathcal{D})} \quad \forall\theta' \tag{A.14}$$

Coming back to the original notations, and performing the integration by parts of the kind $\left(\vec{\Omega}\cdot\nabla u, v\right)_{L_2(\mathcal{D})} = -\left(\vec{\Omega}\cdot\nabla v, u\right)_{L_2(\mathcal{D})} + \left(\vec{\Omega}\cdot\vec{n}u, v\right)_{L_2(\partial\mathcal{D})}$ gives the adjoint problem:

$$\begin{cases} \left[-\vec{\Omega}\cdot\nabla - \dfrac{i\omega}{c} + \kappa + \sigma_s\right] I_s^*(r,\omega) - \dfrac{\sigma}{4\pi}\displaystyle\int_{4\pi} I_s^*\Phi(\vec{\Omega}',\vec{\Omega})\,\mathrm{d}\Omega' = 0 & \forall x\in\mathcal{D} \\ \vec{\Omega}\cdot\vec{n} I_s^* + (P-M)\,P' = 0 & \forall x\in\partial\mathcal{D}^- \\ \left[-\vec{\Omega}\cdot\nabla - \dfrac{i\omega}{c} + \kappa + \sigma_s\right] I_c^*(r,\omega) - \dfrac{\sigma}{4\pi}\displaystyle\int_{4\pi} I_s^*\delta(\vec{\Omega}'-\vec{\Omega}_c)\Phi(\vec{\Omega}',\vec{\Omega})\,\mathrm{d}\Omega' = 0 & \forall x\in\mathcal{D} \\ I_c^* = 0 & \forall x\in\partial\mathcal{D}^- \end{cases} \tag{A.15}$$

Making the distinction for both component types within the cost differential:

$$j'(\theta;\kappa') = \langle \kappa' I_c, I_c^*\rangle_{L_2(\mathcal{D})} + \langle \kappa' I_s, I_s^*\rangle_{L_2(\mathcal{D})} + \{\epsilon(\kappa-\kappa_0),\kappa'\}_{\mathcal{U}} \tag{A.16}$$

$$\begin{aligned} j'(\theta;\sigma') = {} & \langle \sigma' I_c, I_c^*\rangle_{L_2(\mathcal{D})} + \langle \sigma' I_s, I_s^*\rangle_{L_2(\mathcal{D})} + \{\epsilon(\sigma-\sigma_0),\sigma'\}_{\mathcal{U}} \\ & - \left\langle \frac{\sigma'}{4\pi}\int_{4\pi}\left[I_s(r,\vec{\Omega}',\omega) + I_c(r,\omega)\delta(\vec{\Omega}'-\vec{\Omega}_c)\right]\Phi(\vec{\Omega}',\vec{\Omega})\,\mathrm{d}\Omega', I_s^* \right\rangle_{L_2(\mathcal{D})} \end{aligned} \tag{A.17}$$

with null boundary conditions for both I_s^* and I_c^* as the boundary condition of the forward model does not depend on any optical parameter.

The inner products involved in $(\nabla j(\theta), \theta') = j'(\theta;\theta')$ used to extract the cost function gradient from (A.16)-(A.17) is to be defined. The most usual definitions for this inner product is the $L_2(\mathcal{D})$ product, that is $(u,v) := \langle u, v\rangle_{L_2(\mathcal{D})}$ as defined in (A.10), but some alternative norms may also be considered as will be presented in next section. With such inner products definition, i.e. $j'(\theta;\theta') = \langle \nabla j, \theta'\rangle_{L_2(\mathcal{D})}$, one obtains straight-forwardly:

$$\nabla_\kappa j(\theta) = \overline{I_c} I_c^* + \overline{I_s} I_s^* + \epsilon(\kappa - \kappa_0) \tag{A.18}$$

$$\begin{aligned} \nabla_\sigma j(\theta) &= \overline{I_c} I_c^* + \overline{I_s} I_s^* + \epsilon(\sigma - \sigma_0) \\ &- \frac{\overline{I_s^*}}{4\pi} \int_{4\pi} \left[I_s(r, \vec{\Omega}', \omega) + I_c(r,\omega)\delta(\vec{\Omega}' - \vec{\Omega}_c) \right] \Phi(\vec{\Omega}', \vec{\Omega})\, \mathrm{d}\Omega' \end{aligned} \tag{A.19}$$

Such an inner product definition leads to consider, when the adjoint states are known, the cost function gradients simply with a sum of products (see the right hand side of (A.18)-(A.19)) where the adjoint states are given through the integration of the coupled system (A.15).

References

[1] A. Charette, J. Boulanger, and H. K Kim. An overview on recent radiation transport algorithm development for optical tomography imaging. Journal of Quantitative Spectroscopy and Radiative Transfer, 109(17-18):2743–2766, November-December 2008.

[2] Gabor T. Herman. Fundamentals of Computerized Tomography: Image Reconstruction from Projections. Springer, 2009.

[3] S R Arridge. Optical tomography in medical imaging. Inverse Problems, 15(2):R41–R93, 1999.

[4] H. K. Kim and A. Charette. Frequency domain optical tomography using a conjugate gradient method without line search. Journal of Quantitative Spectroscopy and Radiative Transfer, 104(2):248–256, 2007.

[5] J. Boulanger and A. Charette. Reconstruction optical tomography using transient radiative transfert equation and puls laser: a numerical study. Journal of Quantitative Spectroscopy & Radiative Transfer, 93:325–336, 2005.

[6] Alexander Klose. Optical tomography based on the Equation of Radiative Transfer. PhD thesis, Department of Physiscs Freie Universität, Berlin Germany, October 2001.

[7] K. Ren, G. Ball, and A. H. Hielscher. Frequency domain optical tomography based on the equation of radiative transfer. Society for Industrial and Applied Mathematics, 28(4):1463–1489, 2006.

[8] X. Ciu and Ben Q. Li. Discontinuous finite element solution of 2-d radiative transfer with and without axisymmetry. Journal of Quantitative Spectroscopy & Radiative Transfer, 96:383–407, 2005.

[9] H. Gao and H. Zhao. A fast-forward solver of radiative transfer equation. Transport Theory and Statistical Physics, 38:149–192, 2009.

[10] O. Balima, Y. Favennec, J. Boulanger, and A. Charette. Optical tomography with the discontinuous galerkin formulation of the radiative transfer equation in frequency domain. Journal of Quantitative Spectroscopy and Radiative Transfer, 113(10):805–814, 2012.

[11] Ben Q. Li. Discontinuous Finite Elements in Fluid Dynamics and Heat Transfer. Springer-Verlag, 2006.

[12] T. J. R. Hughes, G. Engel, L. Mazzei, and M.G. Larson. A comparison of discontinuous and continuous Galerkin methods based on error estimates, conservation, robustness and efficiency. Springer-Verlag, 2000.

[13] T. J. Chung. Computational fluid dynamics. Cambridge University press, 2010.

[14] K. Levenberg. A method for the solution of certain nonlinear problems in least squares. Quarterly Journal of Applied Mathmatics, II(2):164–168, 1944.

[15] C. Hansen and D.P. O'Leary. The use of the L-curve in the regularization of discrete ill-posed problems. Siam Journal of Scientific Computing, 14(6):1487–1503, 1993.

[16] J. W. Neuberger. Steepest descent and differential equations. Journal of the Mathematical Society of Japan, 37:187–195, 1985.

[17] Max D. Gunzburger. Perspectives in flow control and optimization. SIAM, Philadelphia, 2003.

[18] A. Klose. The forward and inverse problem in tissue optics based on the radiative transfer equation: A brief review. Journal of Quantitative Spectroscopy & Radiative Transfer, 111:1852–1853, 2010.

[19] Michael F. Modest. Radiative Heat Transfer. Academic Press, second edition, 2003.

[20] O. Balima, A. Charette, and D. Marceau. Comparison of light transport models in view of optical tomography applications. Journal of Computational and Applied Mathematics, 234(7):2259–2271, August 2010.

[21] J. Boulanger and A. Charette. Numerical developments for short-pulsed near infra-red laser spectroscopy. part ii: Inverse treatment. Journal of Quantitative Spectroscopy & Radiative Transfer, 91(3):297–318, 2005.

[22] O. Balima, J. Boulanger, A. Charette, and D. Marceau. New developments in frequency domain optical tomography. Part I. forward model and gradient computation. Journal of Quantitative Spectroscopy and Radiative Transfer, 112(7):1229–1234, May 2011.

[23] O. Balima, J. Boulanger, A. Charette, and D. Marceau. New developments in frequency domain optical tomography. Part

II. application with a l-bfgs associated to an inexact line search. Journal of Quantitative Spectroscopy and Radiative Transfer, 112(7):1235–1240, May 2011.

[24] T. Tarvaimen, M. Vauhkonen, and S R Arridge. Gauss–newton reconstruction method for optical tomography using the finite element solution of the radiative transfer equation. Journal of Quantitative Spectroscopy and Radiative Transfer, 109(17-18):2767–2778, 2008.

[25] O. Balima, Y. Favennec, J. Boulanger, and A. Charette. A modified cost function-based optical tomography reconstruction algorithm optimized for finite-elements methods. In ASME 2011 International Mechanical Engineering Congress & Exposition, 11-17 November, Denver, Colorado 2011 (accepted number IMECE2011-63820).

[26] A.N. Tikhonov and V.Y. Arsenin. Solution of ill-posed problems. John Wiley, New-York, 1977.

[27] M. Minoux. Programmation mathématique : théorie et algorithms, tomes 1 et 2. Dunod, Paris, 1983.

[28] M. Minoux. Mathematical Programming, Theory and Applications. Wiley, Chichester, UK, 1986.

[29] P.E. Gill, W. Murray, and M.H. Wright. Practical Optimization. Academic Press, London, 1992.

[30] J. Nocedal and S.J. Wright. Numerical Optimization. Springer, New York, 1999.

[31] Y. Favennec, V. Labbé, and F. Bay. The ultraweak time coupling in nonlinear multiphysics modeling and realated optimization problems. International Journal for Numerical Methods in Engineering, 60(4):803–831, 2004.

[32] O. Balima, Y. Favennec, D. Petit, and A. Charette. Reduced modelling through identification on 2d incompressible laminar flows. Inverse Problems in Science and Engineering, 17(3):303–319, March 2009.

[33] Y. Favennec, V. Labbé, and F. Bay. Induction heating processes optimization – a general optimal control approach. Journal of Computational Physics, 187:68–94, 2003.

[34] Y. Rouizi, Y. Favennec, J. Ventura, and D. Petit. Numerical model reduction of 2D steady incompressible laminar flows: Application on the flow over a backward-facing step. Journal of Computational Physics, 228(6):2239–2255, 2009.

[35] Y. Favennec, P. Le Masson, and Y. Jarny. Lecture 7: Optimization methods for non linear estimation and function estimation. Eurotherm Spring School METTI 2011 : Thermal measurements and inverse techniques: http://www.sft.asso.fr, Roscoff, June 2011.

[36] D. W. Marquardt. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. SIAM Journal on Applied Mathematics, 11(2):431–441, 1963.

[37] V.A. Morozov. Methods for solving incorrectly posed problems. Springer, 1984.

[38] H. Jiang. Diffuse Optical Tomography - Principles and applications. CRC Press, Boca Raton London New York, 2011.

[39] T. Tarvainen, V. Kolehmainen, S.R. Arridge, and J.P. Kaipio. Image reconstruction in diffuse optical tomography using the coupled radiative transport-diffusion model. Journal of Quantitative Spectroscopy & Radiative Transfer, 112:2600–2608, 2011.

[40] D.C. Liu and J. Nocedal. On the limited memory BFGS method for large scale optimization. Mathematical Programming: Series A and B, 45(3):503–528, December 1989.

[41] Y. Jarny. Lecture 9: Inverse problems & regularized solutions. Eurotherm Spring School METTI 2011 : Thermal measurements and inverse techniques: http://www.sft.asso.fr, Roscoff, June 2011.

[42] O. Balima, Y. Favennec, and D. Petit. Model reduction on heat conduction with radiative boundary conditions using the modal identification method. Numerical Heat Transfer Part B, 52(2):107–130, 2007.

[43] G. Chavent. Nonlinear least squares for inverse problems. Springer, 2009.

[44] K.A. Woodbury. Inverse Engineering Handbook. CRC Press, Boca Raton, 2003.

[45] Huabei Jiang. Difuse optical tomography : Principles and Applications. Taylor and Francis, 2011.

[46] P. Destuynder. Analyse et traitement des images numériques. Hermès–Lavoisier, 2006.

[47] R.J. Renka. Image segmentation with a Sobolev gradient method. Nonlinear Analysis, 71:774–780, 2009.

[48] I. Danaila and P. Kazemi. A new Sobolev gradient method for direct minimization of the gross-pitaevskii energy with

rotation. Siam Journal of Scientific Computing, 32(5):2447–2467, 2010.

[49] A. Majid and S. Sial. Application of Sobolev gradient method to Poisson–Boltzmann system. Journal of Computational Physics, 229:5742–5754, 2010.

[50] N. Raza, S. Sial, and S. Siddiqi. Sobolev gradient approach for the time evolution related to energy minimization of Ginzburg–Landau functionals. Journal of Computational Physics, 228:2566–2571, 2009.

[51] N. Raza, S. Sial, and S. Siddiqi. Approximating time evolution related to Ginzburg–Landau functionals via Sobolev gradient methods in a finite-element setting. Journal of Computational Physics, 229:1621–1625, 2010.

[52] R.J. Renka. A simple explanation of the Sobolev gradient method. Technical report, Computer Science and Engineering: University of North Texas, 2006. www.cse.unt.edu/∼ renka/papers/sobolev.pdf.

[53] S. Sial, J. Neuberger, T. Lookman, and A. Saxena. Energy minimization using Sobolev gradients: application to phase separation and ordering. Journal of Computational Physics, 189:88–97, 2003.

[54] F. Bourquin and A. Nassiopoulos. Inverse reconstruction of initial and boundary conditions of a heat transfer problem with accurate final state. International Journal of Heat and Mass Transfer, 54:3749–3760, 2011.

[55] B. Protas, T. Bewley, and G. Hagen. A computational framework for the regularization of adjoint analysis in multiscale PDE systems. Journal of Computational Physics, 195(1):49–89, 2004.

[56] B. Protas. Adjoint-based optimization of PDE systems with alternative gradients. Journal of Computational Physics, 227:6490–6510, 2008.

[57] A. Cardona and S. Idelsohn. Solution of non-linear thermal transient problems by a reduction method. International Journal for Numerical Methods in Engineering, 23:1023–1042, 1986.

[58] Adrian Doicu, Thomas Trautman, and Franz Schreier. Numerical Regularization for Atmospheric Inverse Problems. Springer, 2010.

[59] Robert J. Renka. Nonlinear least squares and Sobolev gradients. Applied Numerical Mathematics, 65:91–104, 2013.

[60] J. Céa. Optimisation, théorie et algorithmes. Dunod, 1971.

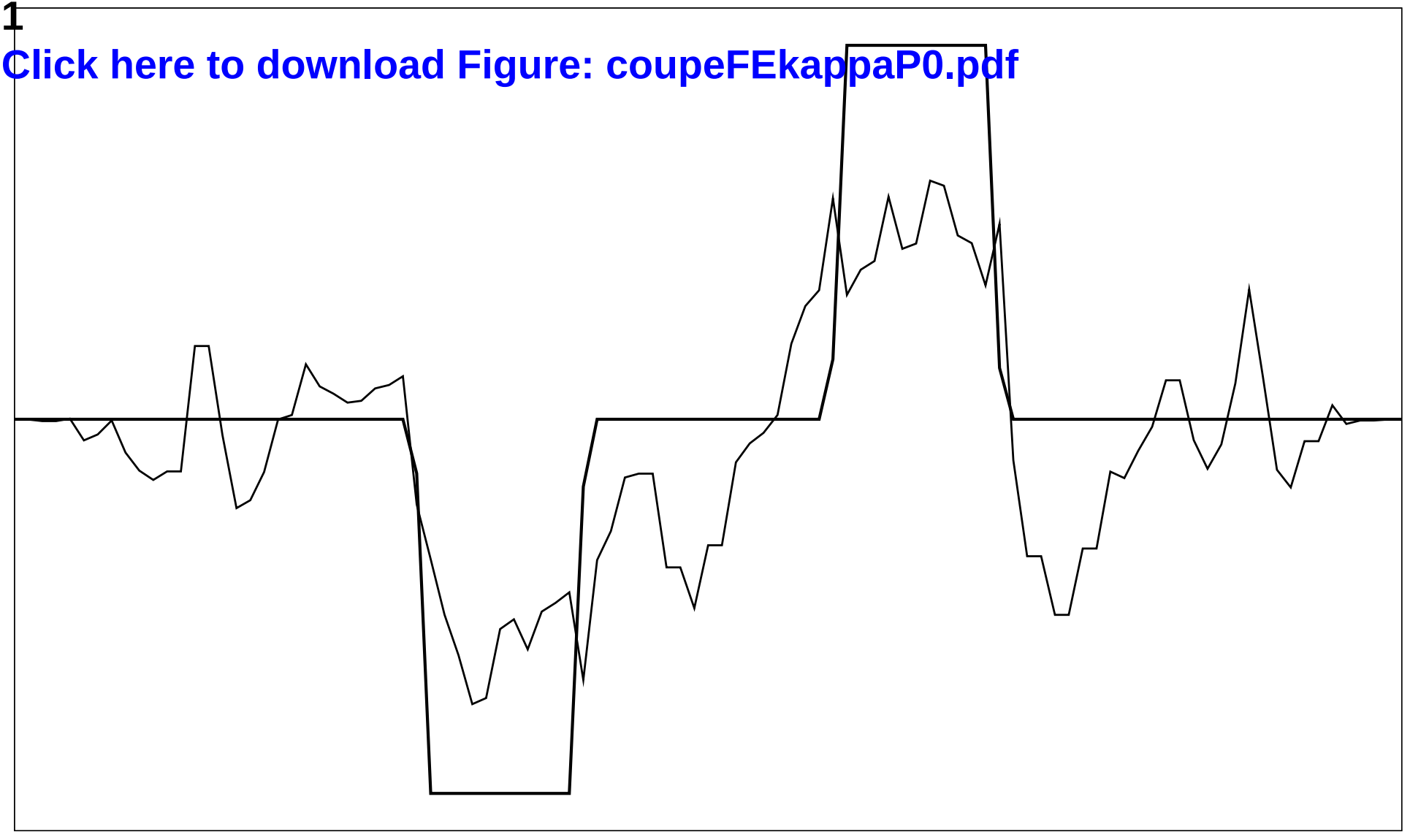
[61] J.-L. Lions. Contrôle optimal des systèmes gouvernés par des équations aux dérivées partielles. Dunod, Paris, 1968.

[62] G. Allaire. Numerical analysis and optimization. Oxford Science Publications, 2007.

[63] P. Michaleris, D.A. Tortorelli, and C. Vidal. Tangent operator and design sensitivity formulations for transient non-linear coupled problems with applications to elastoplasticity. International Journal for Numerical Methods in Engineering, 37:2471–2499, 1994.

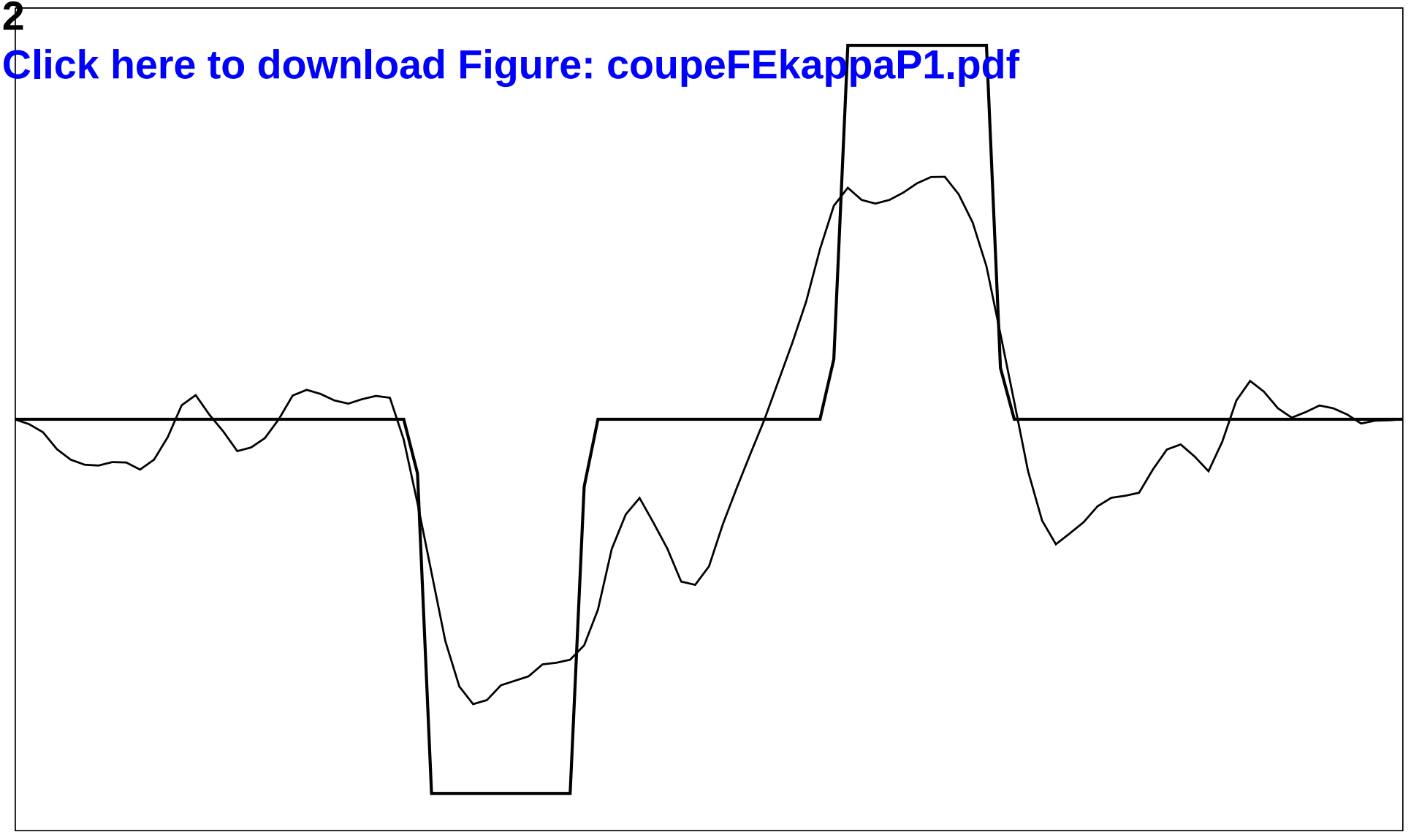
1

Click here to download Figure: coupeFEkappaP0.pdf



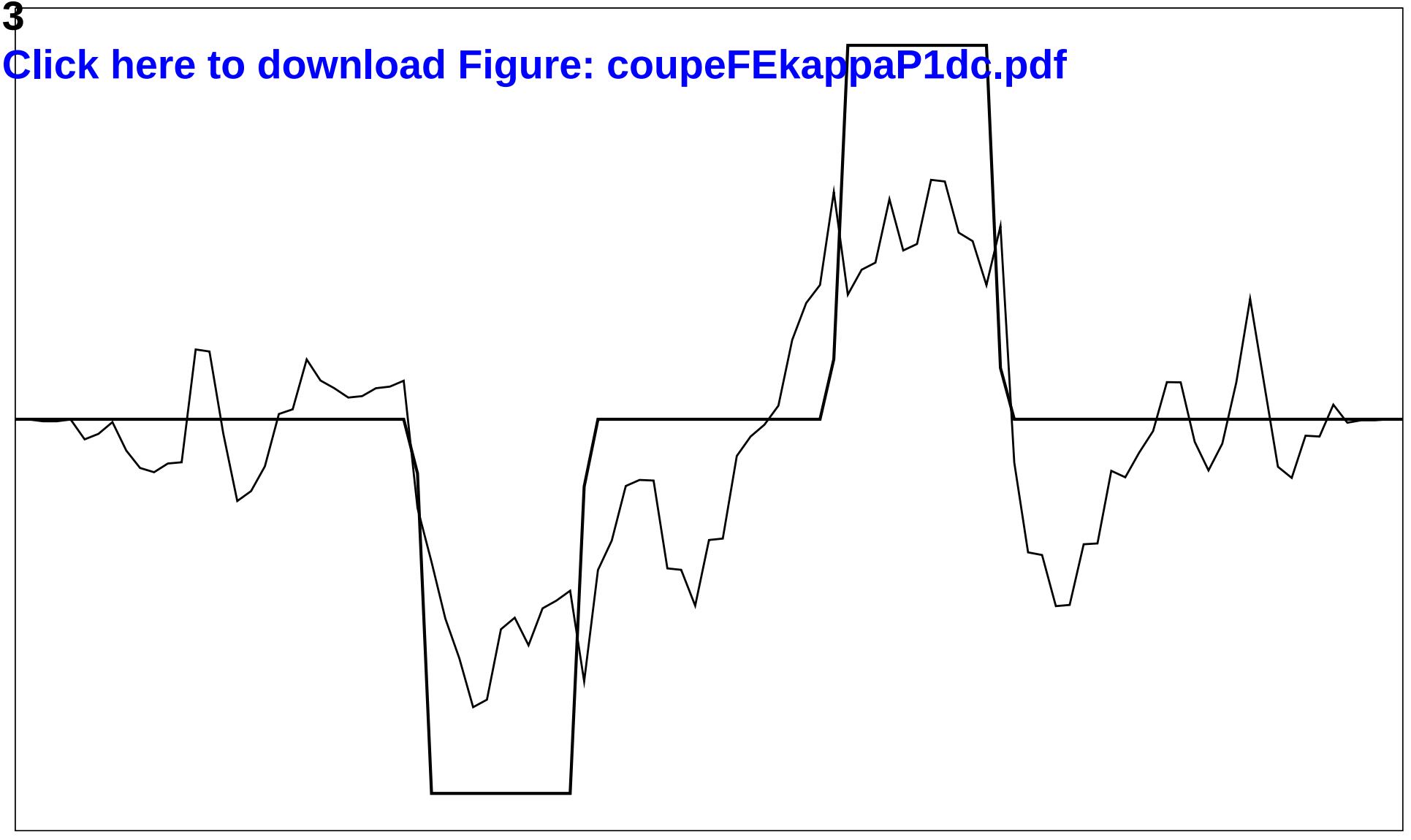
2

Click here to download Figure: coupeFEkappaP1.pdf

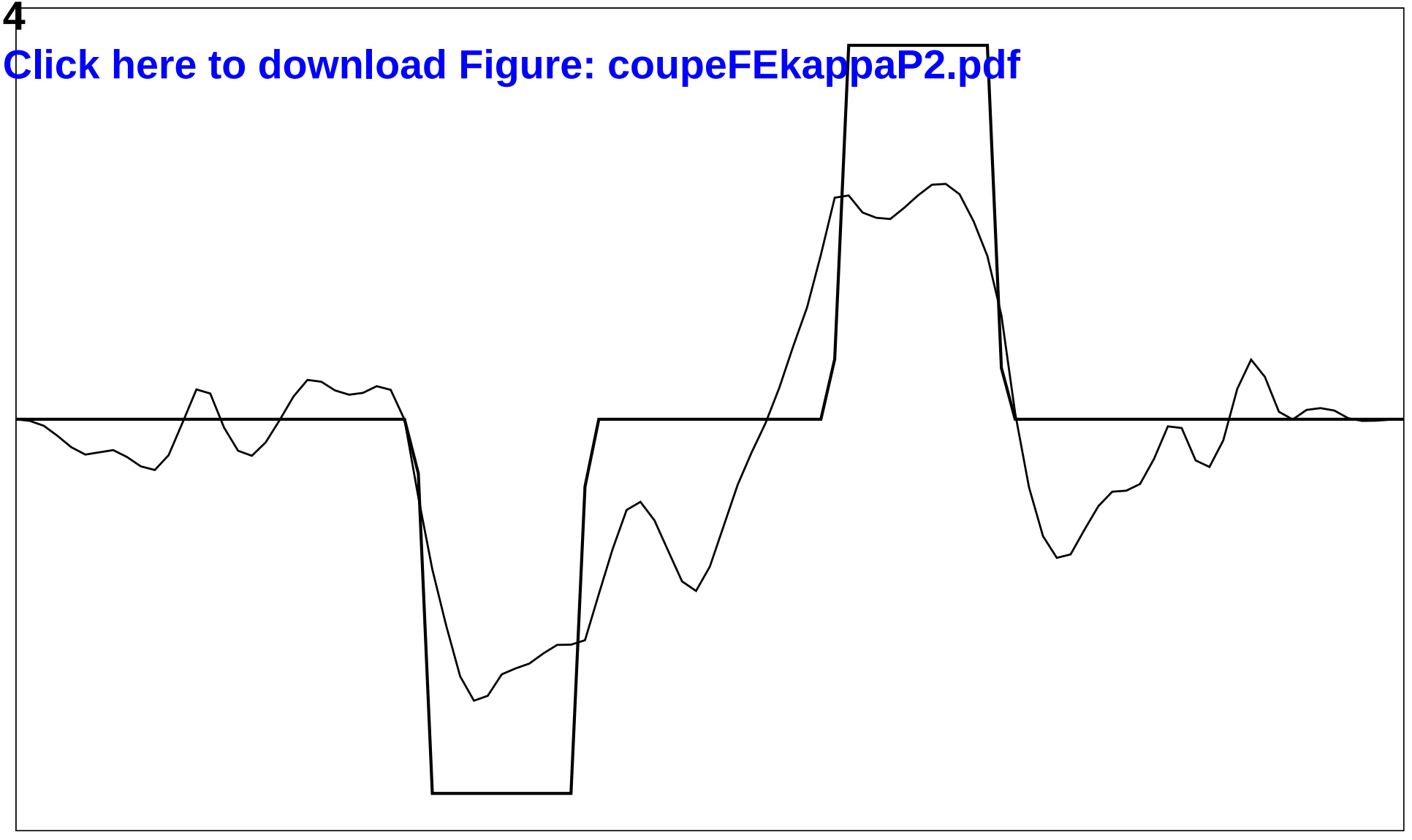


3

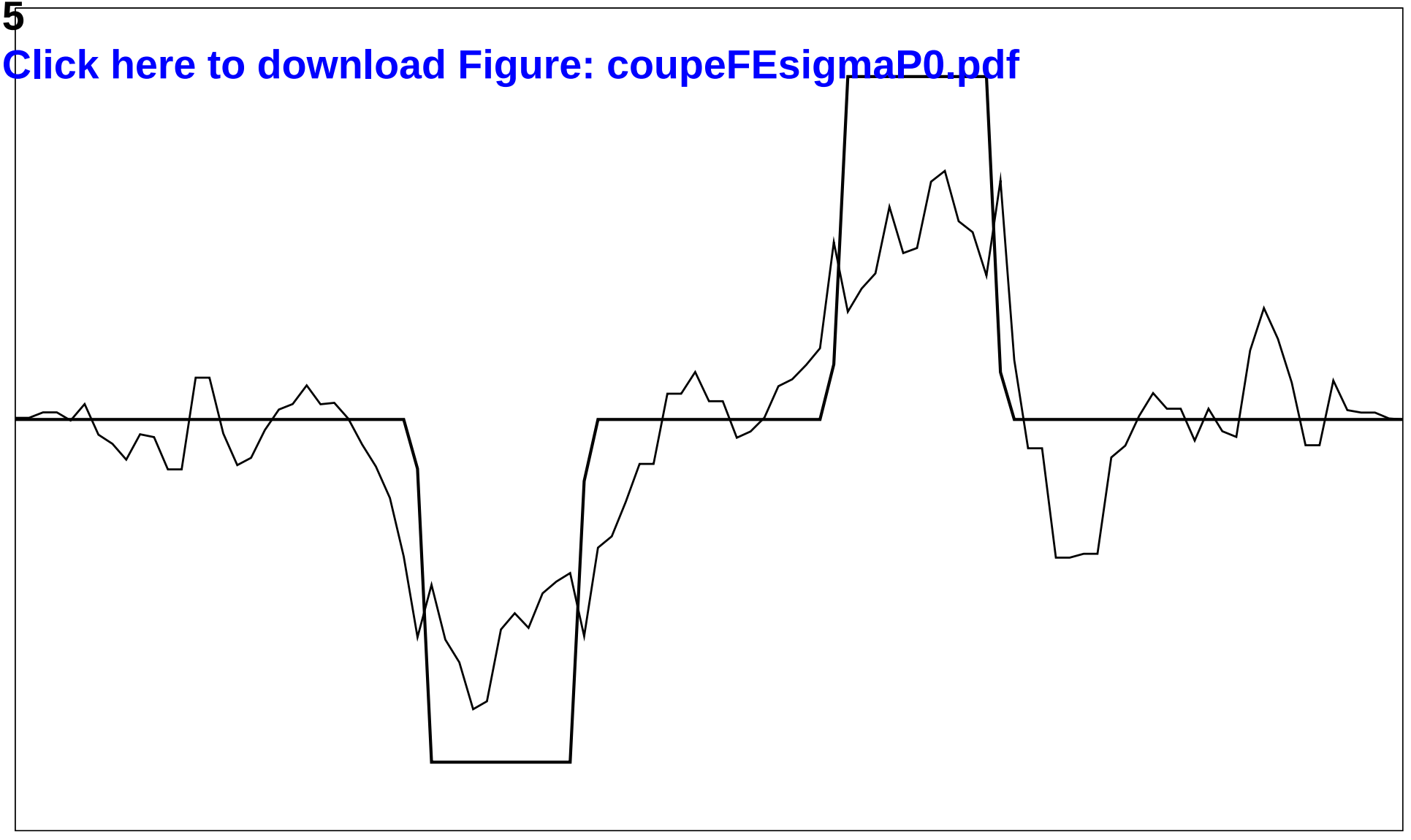
Click here to download Figure: coupeFEkappaP1dc.pdf



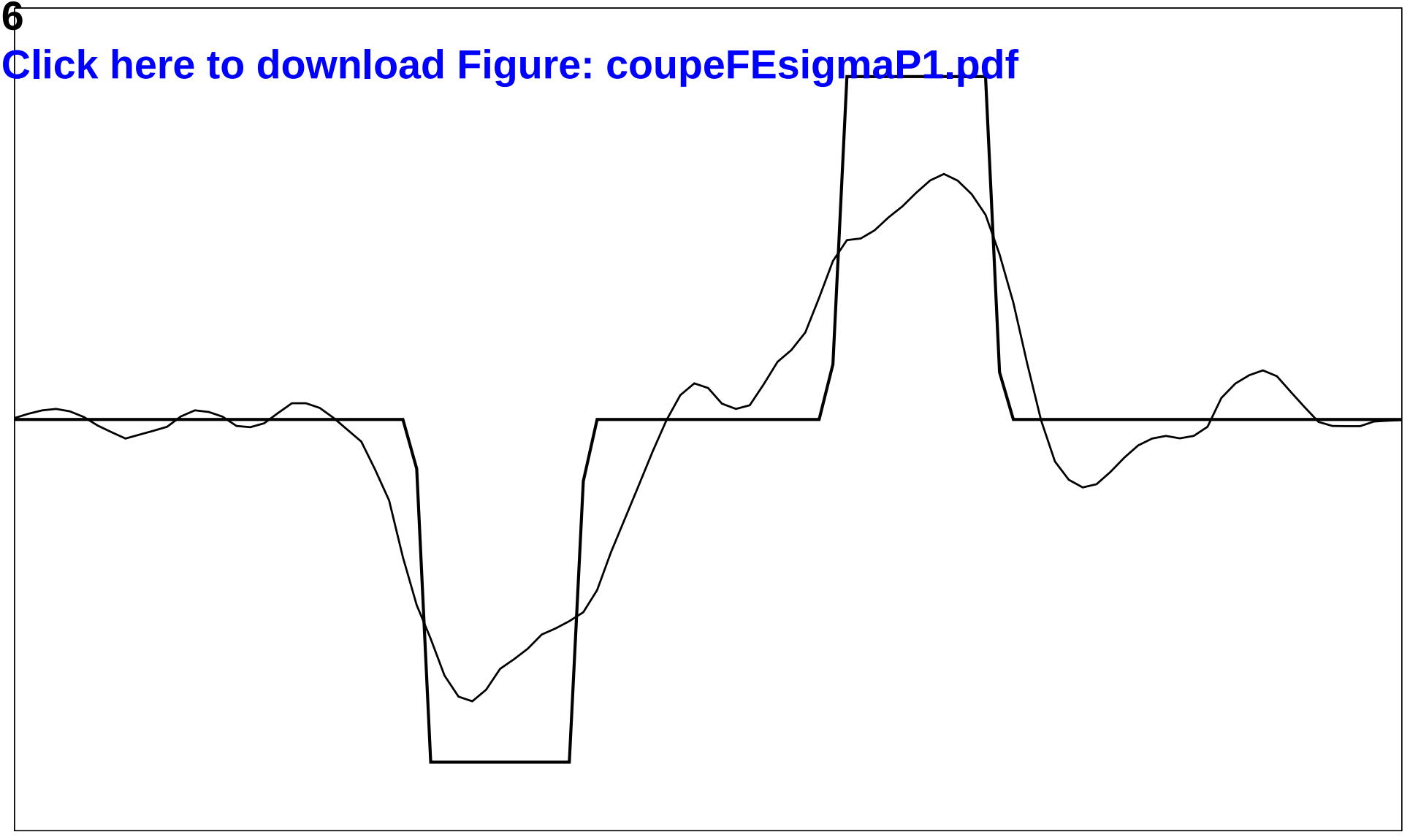
Click here to download Figure: coupeFEkappaP2.pdf



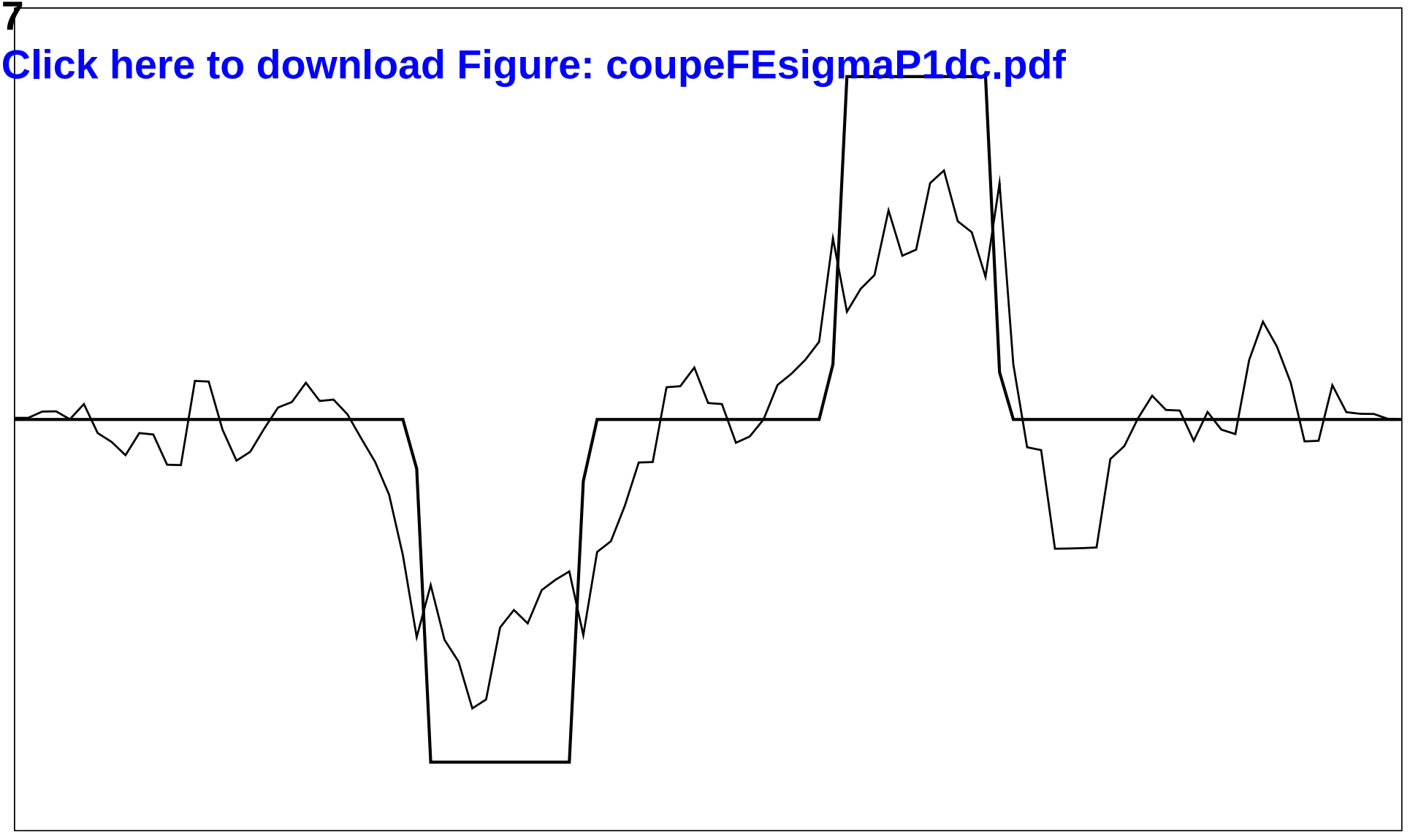
Click here to download Figure: coupeFEsigmaP0.pdf



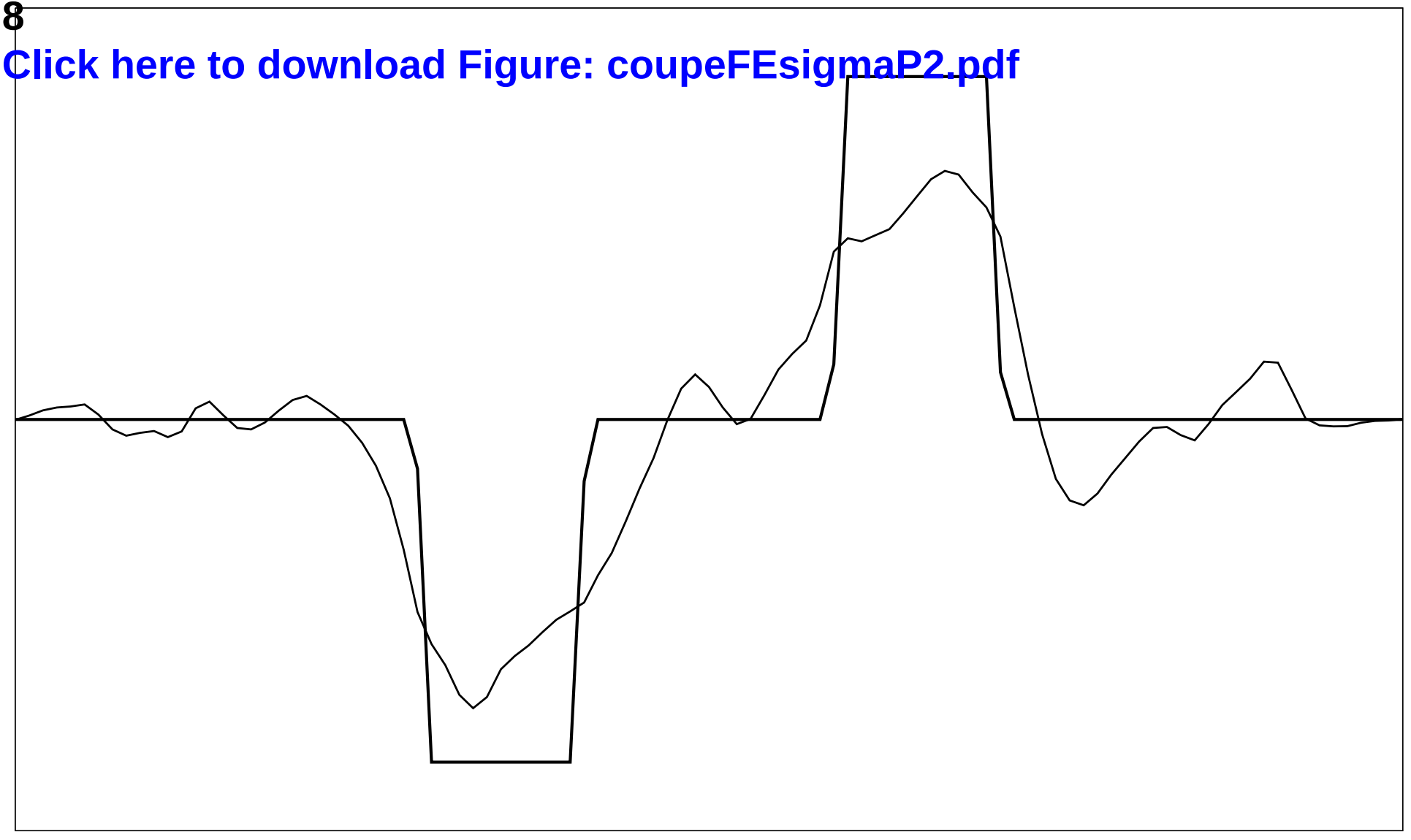
Click here to download Figure: coupeFEsigmaP1.pdf



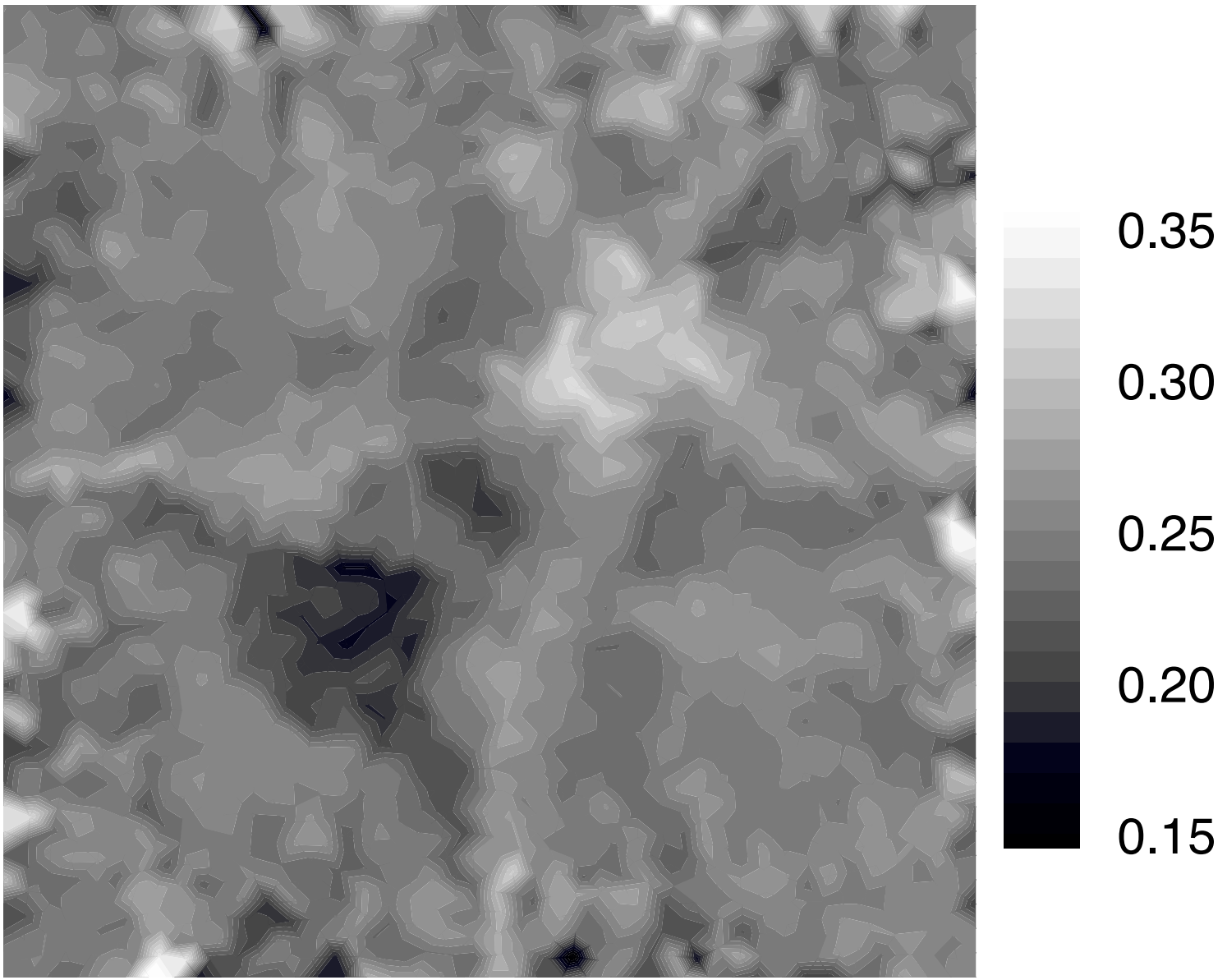
Click here to download Figure: coupeFEsigmaP1dc.pdf



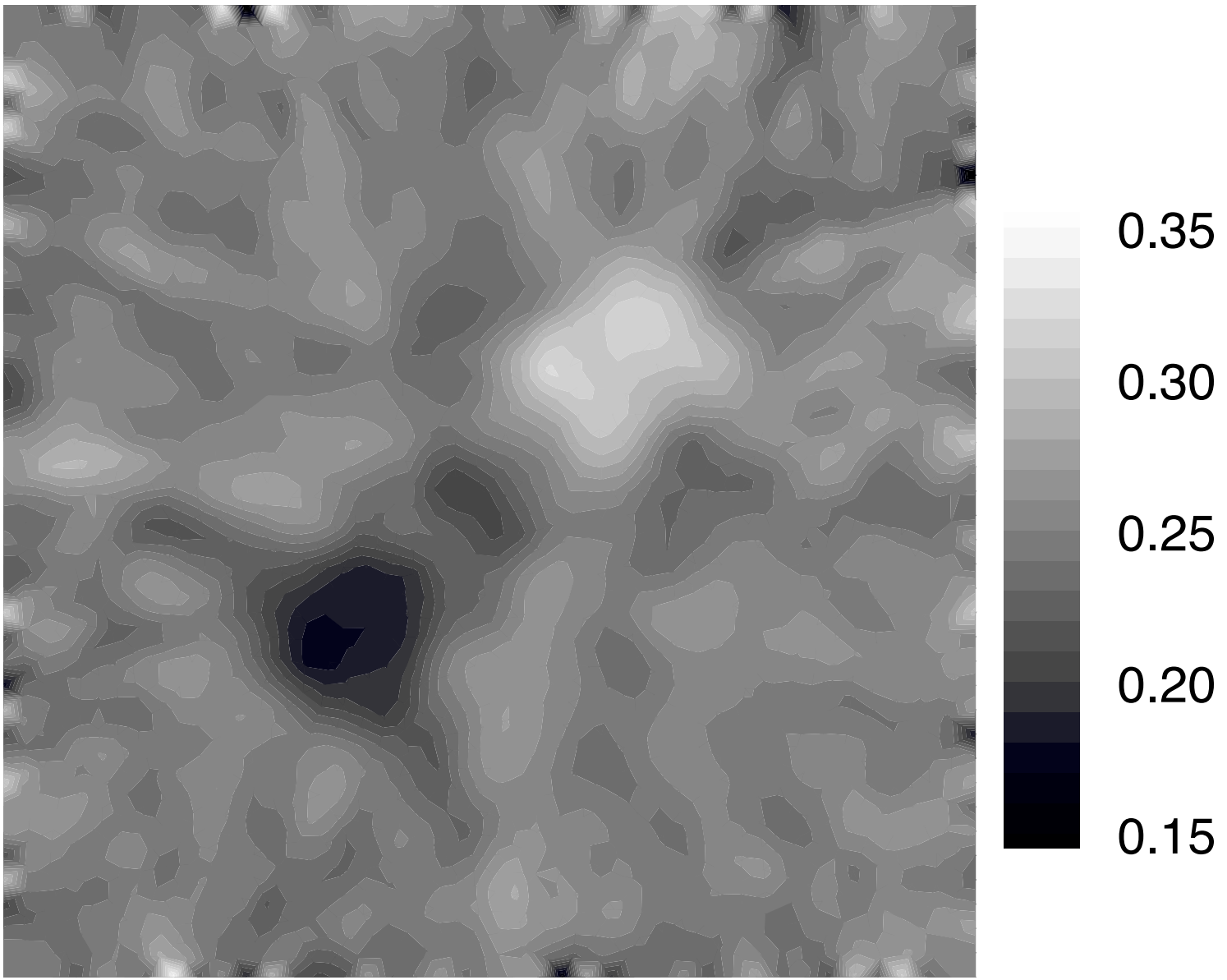
Click here to download Figure: coupeFEsigmaP2.pdf



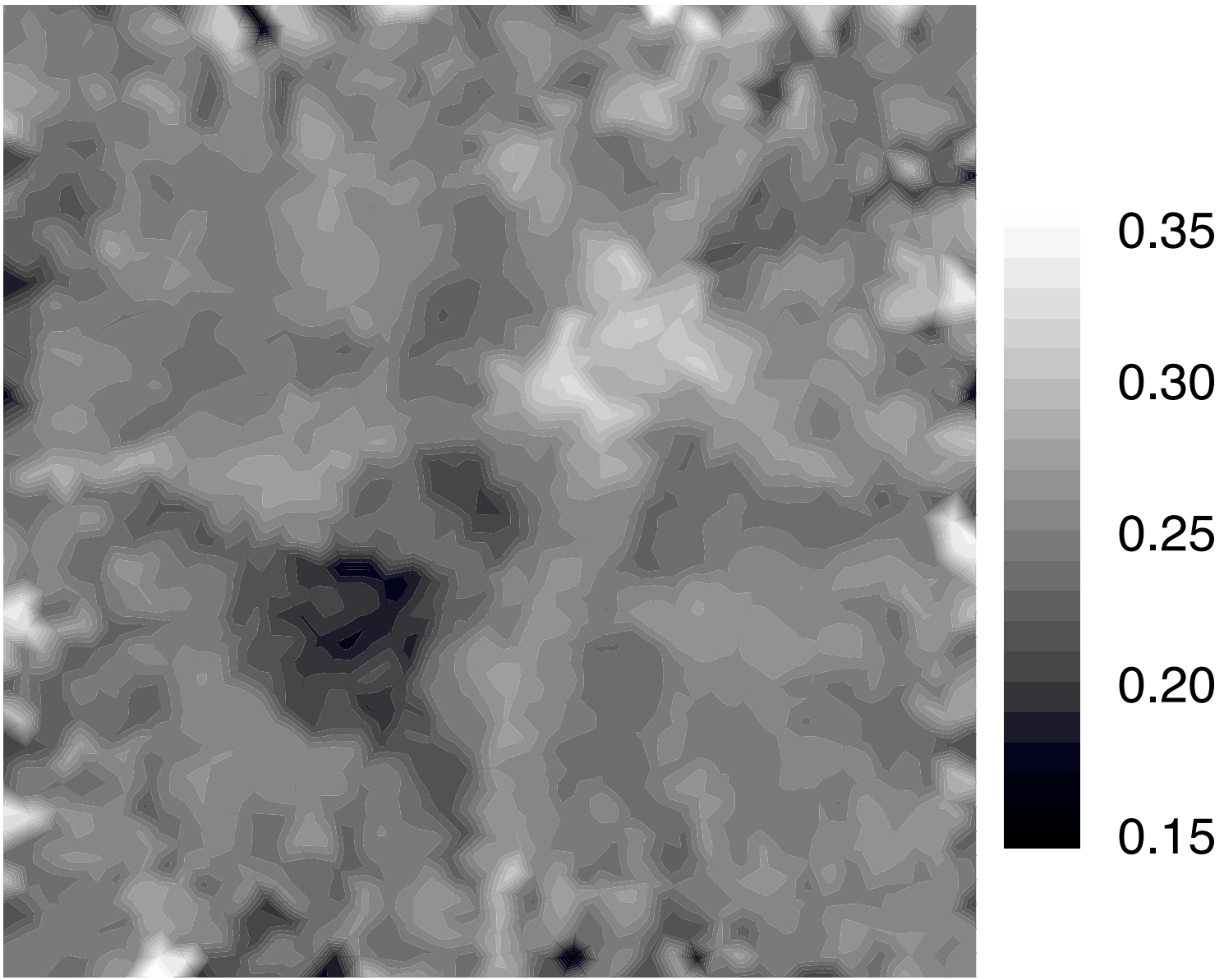
Click here to download Figure: kappaP0.pdf



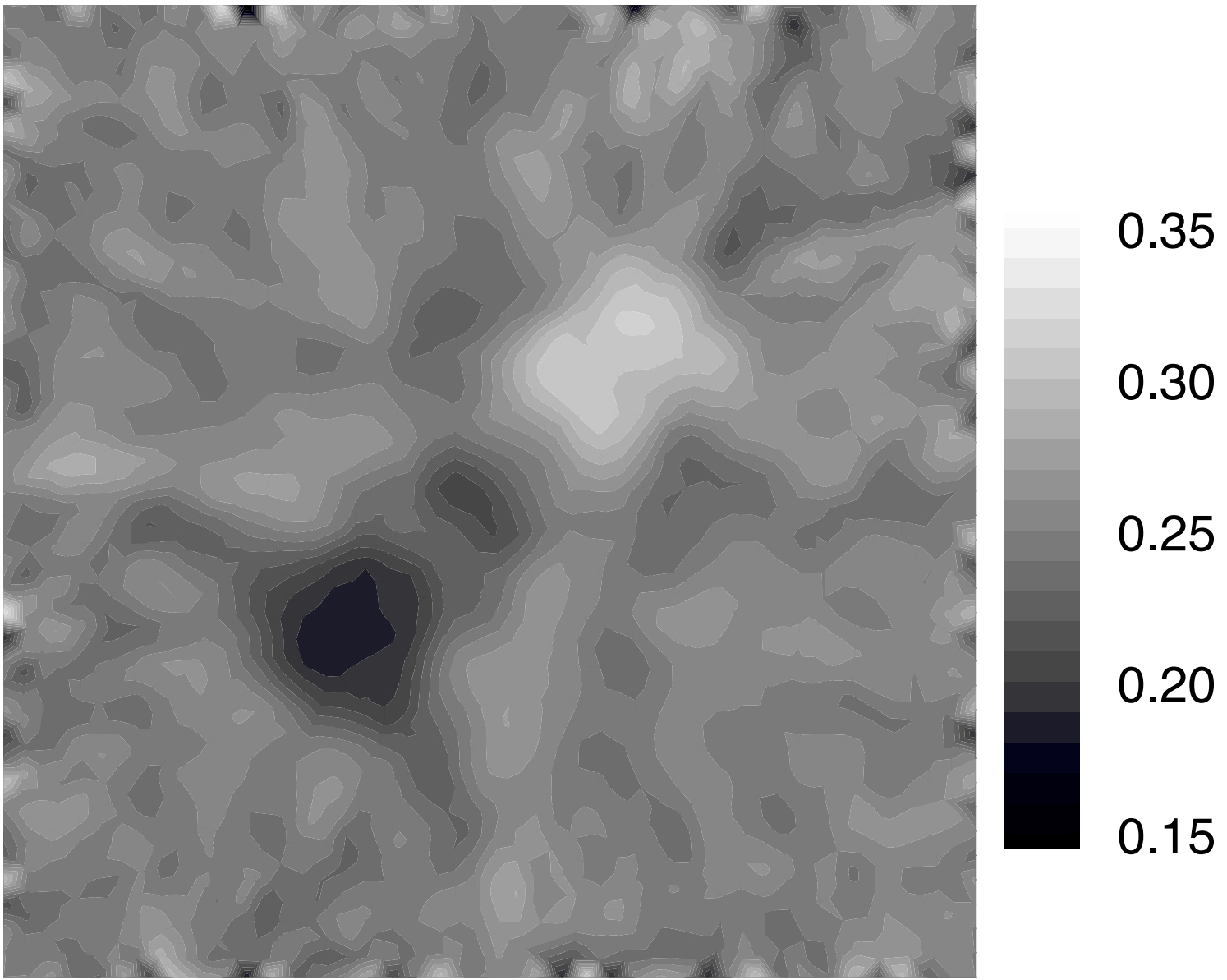
Click here to download Figure: kappaP1.pdf



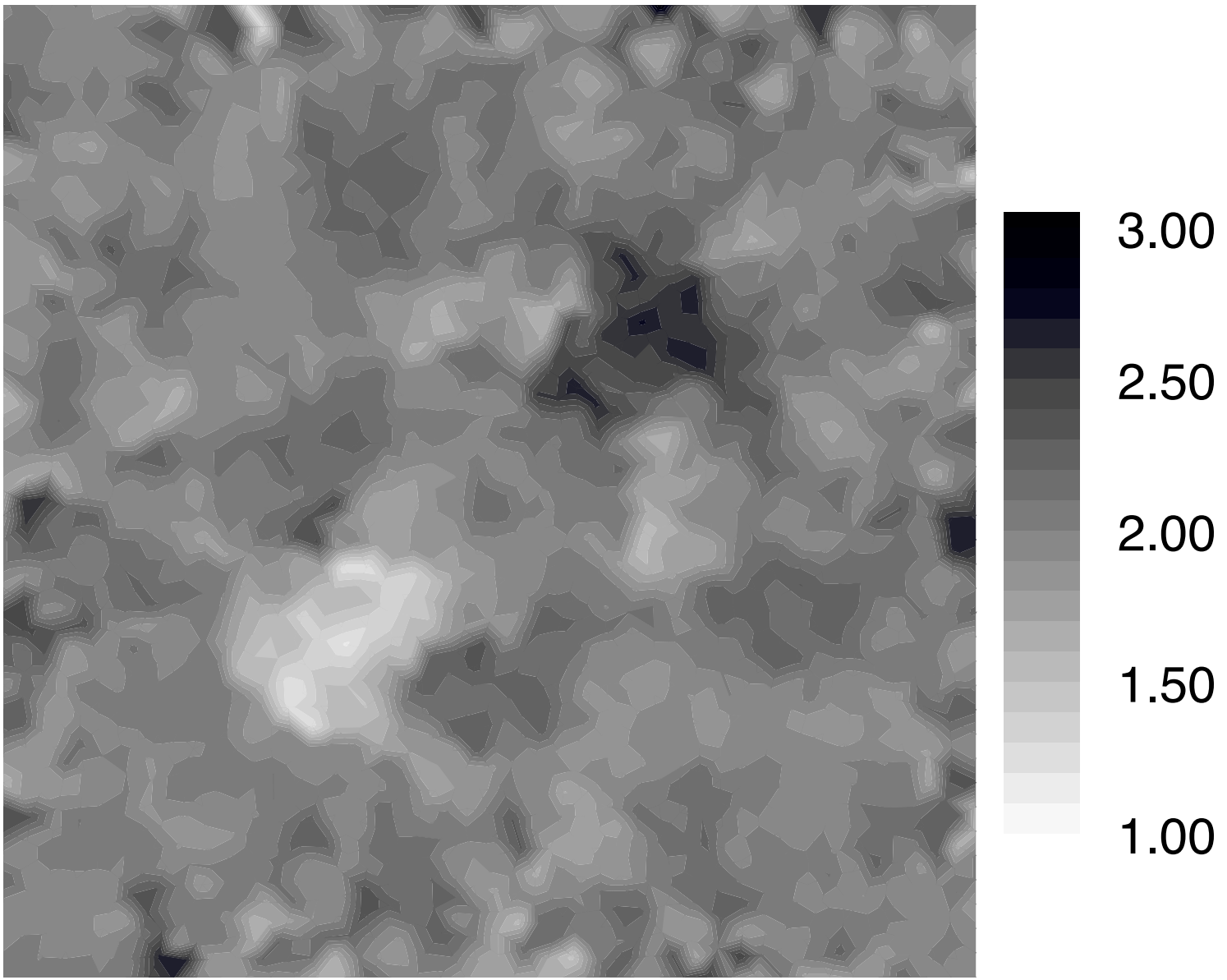
Click here to download Figure: kappaP1dc.pdf



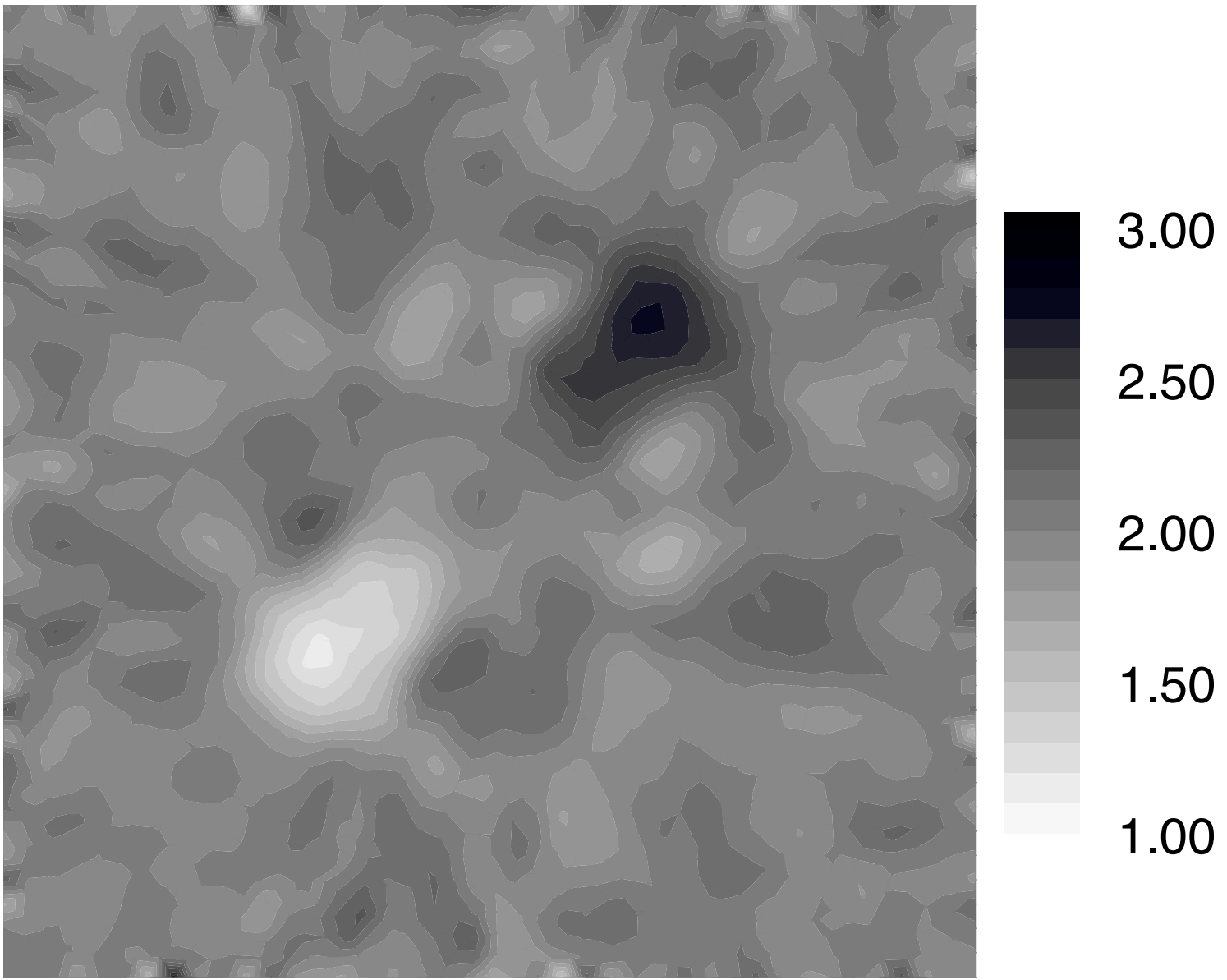
Click here to download Figure: kappaP2.pdf



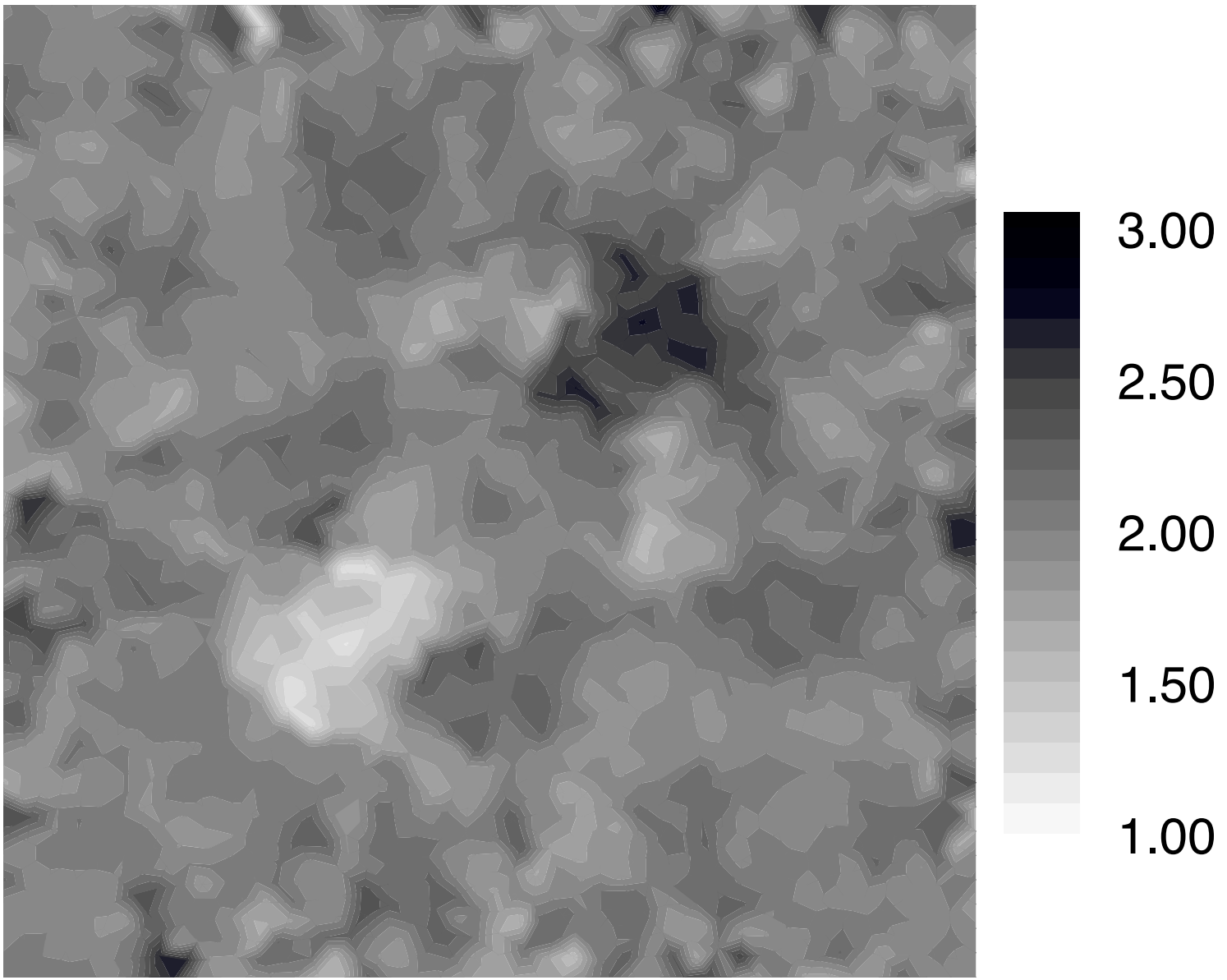
Click here to download Figure: sigmaP0.pdf



Click here to download Figure: sigmaP1.pdf



Click here to download Figure: sigmaP1dc.pdf



Click here to download Figure: sigmaP2.pdf