

# Quasi-optimal Tikhonov penalization and parameterization coarseness in space-dependent function estimation

February 5, 2015

Fabien Dubot <sup>1,2</sup>, Yann Favennec <sup>1</sup>, Benoit Rousseau <sup>1</sup>, Yvon Jarny <sup>1</sup>, Daniel R. Rousse <sup>2</sup>

<sup>1</sup> Laboratoire de Thermocinétique de Nantes (LTN), UMR CNRS 6607, Rue Christian Pauc, BP 50609, 44306 Nantes Cedex 3, France

<sup>2</sup> Chaire de recherche industrielle en technologies de l'énergie et en efficacité énergétique (t3e), École de technologie supérieure, 1100, rue Notre-Dame Ouest, Montréal, Canada H3C 1K3

**Keywords:** inverse problem; Tikhonov regularization; Gauss–Newton; finite element parameterization; ill-posed problems

## Abstract

The determination of space-dependent functions from boundary measurements or inner pointwise measurements are ill-posed inverse problems that require regularization tools to be stabilized. Among numerous regularization strategies, the Tikhonov penalization is one of the most used in the field of space-dependent function estimation. Its efficient use relies on the Tikhonov weight parameter value for which search is time consuming although necessary, specially in the field of non-linear inversion. Other strategies, such as appropriate parameterization, have recently proven to be very efficient to cope with the ill-posedness of such problems. This paper shows that the optimal Tikhonov parameter is almost independent of the mesh used to project the functions to be retrieved. As a consequence, this value should be sought using a coarse mesh even though reconstructions could further be done on finer meshes. This conclusion is validated by numerical means.

## 1 Introduction

The reconstruction of space-dependent functions from pointwise measurements belongs to inverse problems known to be difficult to be solved due to their ill-posed character. More specifically, the reconstruction of maps of radiative properties by illuminating a semi-transparent medium with near infrared radiation and measuring emerging radiation is by definition the classical optical tomography inverse problem.

From a mathematical point of view, within the diffuse approximation framework in the frequency domain [1], the geometry ( $\mathcal{D}$ ) being fixed, the total inward flux ( $I$ ) of the amplitude modulated diffuse source at the frequency  $\nu$  being known as well as the speed of light ( $c$ ) in the medium, the knowledge of the space-dependent absorption ( $\kappa$ ) and reduced scattering ( $\sigma$ ) coefficients is sufficient to simulate the photon density distribution ( $\varphi$ ) in each location of the computational domain. Such modeling is the forward problem which is known to be well-posed mathematically.

Contrarily, the reconstruction of physical properties from the knowledge of photon density on sensors constitutes an inverse problem. Unfortunately, such a problem is ill-posed in the sense of Hadamard [11]. Hence, regularization must be used to stabilize the solution. Several regularization tools have been introduced in the field of optical tomography these last decades, among which, the use of appropriate control space parameterization [10, 3, 7, 2]. Dealing with matrix-based inversion, the Tikhonov regularization method, which is a popular method, consists in adding a penalization term to the cost function to be minimized; this method actually relies on a weight parameter that is to be determined very carefully since: (i) under-regularization leads to small cost functions at the end but at the price of highly fluctuating property maps; and (ii) over-regularization leads to stable property maps around priors, leading to biased solutions [18].

In the following, the optimal Tikhonov parameter,  $\vartheta^*$ , is defined as the one that minimizes the distance between the actual solution and the noise-free solution. This distance can be computed with  $E(\vartheta) = \frac{\|\gamma_\vartheta - \bar{\gamma}\|}{\|\bar{\gamma}\|}$ , where  $\vartheta$  is the Tikhonov parameter,  $\gamma_\vartheta$  is the solution given for the Tikhonov parameter  $\vartheta$ ,  $\bar{\gamma}$  is the exact solution, and  $\|\cdot\|$  is a norm. The point is that one cannot, in practice, compute the optimal Tikhonov parameter,  $\bar{\gamma}$  being unknown. However, numerous heuristic methods have been introduced to compute quasi-optimal Tikhonov parameters. Some methods use the assumed to be known standard deviation such as in the discrepancy principle [18, 9]. Others [8, 12, 19] use less information about the noise properties, such as for instance the Generalized Cross Validation (GCV), or even no information about the noise level present in the system, such as the well known and controversial L-curve. The methods used to find the optimal Tikhonov parameter or at least a quasi-optimal parameter is beyond of the scope of this paper. Moreover, even though methods that do not use error informations are widely used for solving practical problems in various engineering areas, let us recall that in general the L-curve method for instance, because it does not rely on any noise information, is, as shown by [26, 25], not convergent and introduces a nonremovable bias.

The remainder of the paper will show that the optimal Tikhonov parameter is almost independent of the mesh coarsening used for the parameterization. This assertion is demonstrated theoretically and verified numerically based on two different under-determined inverse problems of space-dependent function estimation: a linear inverse heat conduction problem of flux estimation and a non-linear diffuse optical tomography problem. The studies are performed on synthetic data and therefore the errors on reconstructions could be calculated. For this reason one can somehow consider that the proposed approach is heuristic.

Before dealing with the non-linear inverse problem of diffuse optical tomography (which

is the main objective of this paper), a steady-state two-dimensional inverse heat conduction problem of space-dependent heat flux estimation is dealt with in section 2. Such inverse problem has recently been studied [14, 22, 6, 15]. This space-dependent heat flux estimation problem is highly different from the diffuse optical tomography problem because: (i) it deals with a different physics: heat conduction; (ii) measurements are performed within the domain rather than on a boundary; (iii) the heat flux on a boundary is the function to be retrieved rather than physical properties within the medium; and (iv) the forward problem is linear with respect to the coefficients involved in the space-dependent finite element heat flux parameterization. The interest of this first study is above all to illustrate that the quasi-independence of the optimal Tikhonov parameter with the control space parameterization is found in a less heuristic way since only measurement errors are used in the whole process. Indeed, the determination of the Tikhonov parameter can be performed straightforwardly through the knowledge of measurement errors only. In this case, the discrepancy principle is the major recipe to compute the parameters involved in the space-dependent heat flux parameterization. It is known that this method gives satisfactory results and other methods found in literature may give a Tikhonov parameter closer to the optimal one [8, 12, 19] but such consideration is out of the scope of this paper. Moreover, the statement of this article is also validated on this linear inverse problem in the sense of the optimal Tikhonov parameter  $\vartheta^*$  previously defined as  $\arg \min E(\vartheta)$ .

Section 3 deals with the nonlinear diffuse optical tomography problem. The forward model of light propagation in a highly diffuse medium is first presented along with the inverse problem settings. First and second order cost function derivatives used afterwards in optimization are then detailed. The finite element parameterization of, on one side, the state and its derivative and, on the other side, the properties to be retrieved, are presented. The optimization is then written down based on the proposed finite dimensional control space. This section then presents the most usual Tikhonov regularization and proves that, under some weak hypothesis, the optimal Tikhonov parameter is quasi-independent of the dimension of the control space. A numerical verification is performed on the non-linear diffuse optical tomography problem based on the theoretical demonstration.

Overall, the conclusion is that, when dealing with non-linear inverse problems demanding heavy computational time, the determination of the Tikhonov parameter, whatever the method chosen among those of [24] for instance, should be preferably built using a highly coarse mesh in order to lower computational time.

## 2 Space-dependent heat flux estimation

### 2.1 Problem statement

In this first study, a linear steady-state two-dimensional inverse heat conduction problem is considered to illustrate the quasi-independence of the Tikhonov parameter with mesh coarsening used for the finite element projection of the quantity of interest. In the present case already studied in different contexts [14, 22, 6, 15], a space-dependent heat flux on a part

of the boundary is to be retrieved from pointwise measurements within a bounded domain  $\mathcal{D} = [0, 1]^2$ . This application is treated first because of its simplicity: as a matter of fact, the response being linear with respect to the input flux [15], the cost function to be minimized is purely quadratic, and specific tools such as the singular value decomposition coupled with the maximum discrepancy principle [18, 9] can be used to compute straightforwardly a quasi-optimal Tikhonov parameter. In this case, one can speak of ‘‘Tikhonov parameter found according to the discrepancy principle’’. It is shown elsewhere [19] that this quasi-optimal Tikhonov parameter is likely to be close to the optimal one.

Homogeneous steady-state heat conduction without source term but with mixed boundary conditions is described such that:

$$\begin{cases} -\Delta T = 0 & (x_1, x_2) \in \mathcal{D} = ]0, 1[ \times ]0, 1[ \\ T = 0 & x_1 = 0 \\ \nabla T \cdot \mathbf{n} = 0 & x_2 = 0 \text{ and } x_2 = 1 \\ -\lambda \nabla T \cdot \mathbf{n} = \varphi & x_1 = 1 \end{cases} \quad (1)$$

where  $T$  is the temperature,  $\lambda$  is the thermal conductivity and  $\varphi$  is a space-dependent heat flux. The forward problem consists in solving eq. (1) for  $T$  assuming  $\lambda$  and  $\varphi(x_2)$  are known. In contrast, the inverse problem consists in estimating the space-dependent heat flux  $\varphi(x_2)$  on the basis of suitable measurement data  $\check{T}^l$  in  $\mathcal{D}$ , under the assumption that the value of  $\lambda$  is known.

Let us consider  $k$  inner pointwise measurements  $(x_1^l, x_2^l)$ ,  $l = 1, \dots, k$ . Discrepancies between predictions  $T^l = T(x_1^l, x_2^l)$  and associated measurements  $\check{T}^l$ ,  $l = 1, \dots, k$ , are integrated to the cost function:

$$j_\vartheta(\varphi) = \mathcal{J}(T) + \mathcal{J}^+(\varphi) = \left\| T - \check{T} \right\|_{\mathbb{R}^k}^2 + \vartheta \|\phi\|_{\mathbb{R}^\Xi}^2 \quad (2)$$

with  $\|a\|_{\mathbb{R}^k}^2 = \sum_{l=1}^k a_l^2$  and  $\|a\|_{\mathbb{R}^\Xi}^2 = \sum_{l=1}^\Xi a_l^2$ .  $\vartheta$  is the Tikhonov parameter to be searched according to the noise level  $\epsilon$ .  $\phi$  is the vector obtained after finite element parameterization of the heat flux  $\varphi$ , that is  $\varphi(x_2) = \sum_{l=1}^\Xi \Theta_l(x_2) \varphi(x_{2l}) = \sum_{l=1}^\Xi \Theta_l(x_2) \phi_l$  where  $(\Theta_l)_{l=1}^\Xi$  is a finite element basis of  $[0, 1]$ . The singular value decomposition of the matrix  $A : \phi \in \mathbb{R}^\Xi \mapsto T \in \mathbb{R}^k$  such that  $A = W \Lambda V^t$ , enables us to rewrite the cost function such that:

$$j_\vartheta(\xi) = \left\| \Lambda \xi - \check{\xi} \right\|_{\mathbb{R}^k}^2 + \vartheta \|\xi\|_{\mathbb{R}^\Xi}^2 \quad (3)$$

where  $\xi = V^t \phi$  and  $\check{\xi} = W^t \check{T}$ . With such a decomposition, the solution of  $\xi_\vartheta^* = \arg \min j_\vartheta(\xi)$  is given in terms of singular values of  $A$ ,  $(\eta_i)_{i=1}^\Xi$ :

$$\xi_{\vartheta i}^* = \frac{\eta_i \check{\xi}_i}{\eta_i^2 + \vartheta} \quad \forall i = 1, \dots, \Xi \quad (4)$$

On the other hand, the discrepancy principle leads to determine  $\xi_\vartheta^*$  such that:

$$\left\| \Lambda \xi_\vartheta^* - \check{\xi} \right\|_{\mathbb{R}^k}^2 = \epsilon^2 \quad (5)$$

where  $\epsilon$  is the sum of variances of noise on all sensors, i.e.  $\epsilon^2 = \sum_{l=1}^k \epsilon_l^2$ . Combining this latter relationship with eq. (4) leads to determine the quasi-optimal tikhonov parameter found according to the discrepancy principle,  $\vartheta_{\text{dp}}^*$  solution of:

$$R(\vartheta_{\text{dp}}^*) = \sum_{i=1}^k \left( \frac{\vartheta_{\text{dp}}^* \check{\xi}_i}{\eta_i^2 + \vartheta_{\text{dp}}^*} \right)^2 - \epsilon^2 = 0 \quad (6)$$

## 2.2 Numerical results

The numerical study shows that the solution of eq. (6) is almost independent of the dimension  $\Xi$  used in the parameterization of the flux. To do so, let the synthetic data be generated with the flux  $\bar{\varphi}(x_2) = \varphi_0 (\sin(\frac{\pi x_2}{2}) - 1)$ ,  $\varphi_0 = -10^4$  W/m<sup>2</sup> and the thermal conductivity  $\lambda = 30$  W/m K. Five pointwise temperature measurements ( $k = 5$ ) are performed at locations  $x_1 = 0.9$ , and  $x_2 = 0.1, 0.3, 0.5, 0.7$  and  $0.9$ . The data is then perturbed according to a Gaussian white noise with variance  $\epsilon_l^2$  ranging from  $10^{-4}$  to  $10^{-1}$  K<sup>2</sup>,  $\forall l = 1, \dots, k$ , yielding to  $\check{T}$ . The finite element method is used to solve the forward problem eq. (1) based on a regular mesh of the bounded domain  $\mathcal{D}$  and a discretization of the temperature  $T$  with Lagrange  $\mathcal{P}_1$  elements. The regular grid associated to  $\mathcal{D}$  is chosen sufficiently fine to ensure that errors due to the approximation method are negligible when compared to measurement errors. The basis functions  $(\Theta_l)_{l=1}^{\Xi}$  used in the parameterization of the unknown flux  $\varphi$  are also continuous first order Lagrange functions, i.e. linear functions per element satisfying  $\Theta_l(x_{2p}) = 1$  if and only if  $p = l$ .

Table 1 presents the quasi-optimal Tikhonov parameter  $\vartheta_{\text{dp}}^*$  for several uniform Lagrange parameterizations with  $\Xi$  ranging from 7 to 25, and several variances of noise in the data, from  $\epsilon_l^2 = 10^{-4}$  to  $\epsilon_l^2 = 10^{-1}$ . It is seen that, for a given level of noise, the Tikhonov parameter  $\vartheta_{\text{dp}}^*$  depends only very slightly on the flux discretization. The small fluctuations that remain may come from all numerical approximations: finite element computations to build the state matrix  $A$ , the singular value decomposition  $W\Lambda V^t = A$  and the numerical optimization when solving the non-linear problem eq. (6),  $\vartheta_{\text{dp}}^* = \arg \min R(\vartheta_{\text{dp}})$ . Moreover, it is worth noting the linear dependency of  $\vartheta_{\text{dp}}^*$  with the noise variance  $\epsilon_l^2$  (Table 1).

Figure 1 presents the distance from the actual solution to the expected target distribution,  $E(\vartheta) = \frac{\|\varphi_{\vartheta} - \bar{\varphi}\|}{\|\bar{\varphi}\|}$ , as a function of the Tikhonov parameter  $\vartheta$ , for the variance noise  $\epsilon_l = 10^{-3}$ ,  $\forall l$ . The norm was chosen to be defined as  $\|\varphi\| = \left( \int_{x_1=1} \varphi^2 dx_2 \right)^{\frac{1}{2}}$ . Results obtained for other noise variances gave similar curves to the one given in Figure 1 and are thus not presented here. Figure 1 shows that, for low Tikhonov parameters  $\vartheta$ , the distance from the solution to the target decreases when the control space dimension  $\Xi$  decreases. On the contrary, for large Tikhonov parameters, the distance from the solution to the target is independent of the control space dimension. In between, there is a plateau, and the minimum point is found to be independent of  $\Xi$ .

In order to illustrate the effect of regularization, different reconstructions are presented in Figure 2. This figure shows how parameterization influences the reconstructions [7, 2]. It

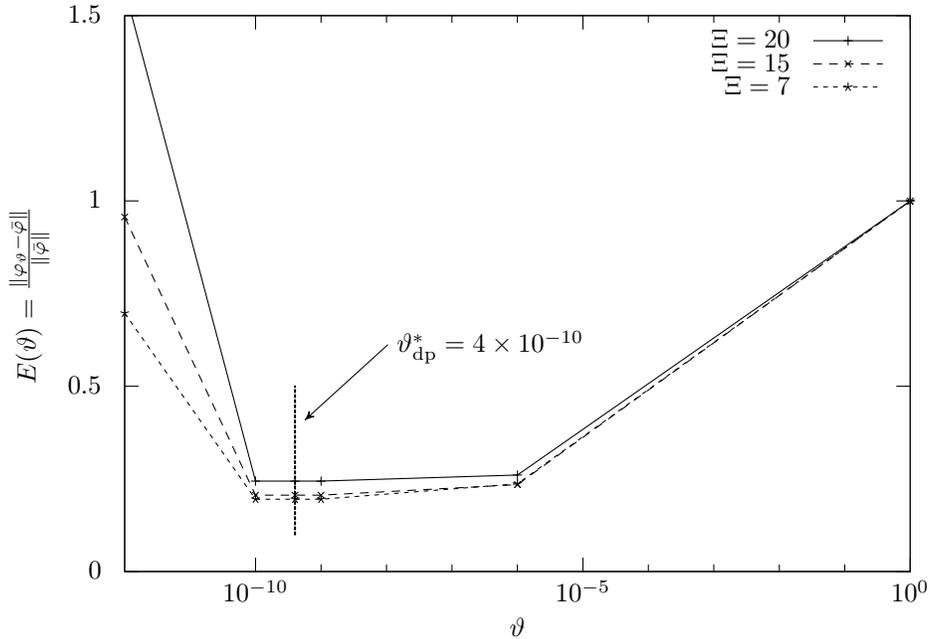


Figure 1: Distance from the solution to the target  $E(\vartheta)$  as a function of the Tikhonov parameter  $\vartheta$ , for three control space dimensions  $\Xi$  equal to 7, 15 and 20. It is seen that  $\vartheta^* = \arg \min E(\vartheta)$  is independent of  $\Xi$ . The quasi-optimal Tikhonov parameter found according to the discrepancy principle has also been added.

is also observed from Figure 2 that despite the very limited number of steady-state measurements (a study on the optimal number and location of inner pointwise measurements is out of the scope of this paper), the space-dependent heat flux could be retrieved for appropriate Tikhonov regularization combined with appropriate control space parameterization.

$\Xi \setminus \epsilon_l^2$	$10^{-4}$	$10^{-3}$	$10^{-2}$	$10^{-1}$
7	$4.189 \times 10^{-11}$	$4.198 \times 10^{-10}$	$4.297 \times 10^{-9}$	$5.493 \times 10^{-8}$
10	$5.246 \times 10^{-11}$	$5.257 \times 10^{-10}$	$5.374 \times 10^{-9}$	$6.622 \times 10^{-8}$
15	$4.240 \times 10^{-11}$	$4.248 \times 10^{-10}$	$4.330 \times 10^{-9}$	$5.151 \times 10^{-8}$
20	$4.042 \times 10^{-11}$	$4.049 \times 10^{-10}$	$4.113 \times 10^{-9}$	$4.683 \times 10^{-8}$

Table 1: Value of the Tikhonov parameter  $\vartheta_{\text{dp}}^*$  solution of (6) for different discretizations and different noise magnitudes.

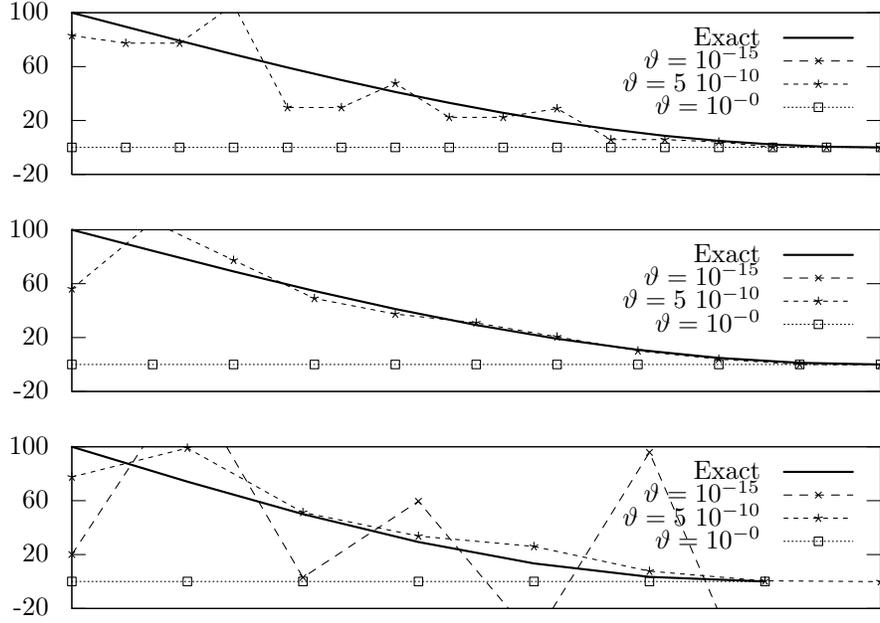


Figure 2: Reconstructions  $\varphi_{\vartheta}(x_2)$  for  $\Xi = 15$  (top),  $\Xi = 10$  (middle) and  $\Xi = 7$  (bottom) for under-regularization ( $\vartheta = 10^{-15}$ ), over-regularization ( $\vartheta = 1$ ) and appropriate regularization ( $\vartheta = 5 \cdot 10^{-10} \approx \vartheta^*$ ). Note that points for  $\Xi = 15$  and  $\Xi = 10$  (over-parameterization) with  $\vartheta = 10^{-15}$  (under-regularization) are not presented because of divergence. The noise variance  $\epsilon_l^2 = 0.001$  was used after generating the synthetic data.

### 3 Optical tomography

#### 3.1 Forward model and inversion setting

Solving for the optical tomography inverse problem is usually based on the minimization of a cost function which depends on the discrepancy between some measurements and the related predictions, the latter being a solution of the forward model which is, in the present case, based on the two-dimensional diffuse approximation. The complex photon density,  $\varphi : \mathcal{D} \mapsto \mathbb{C}$ , is mathematically described by the diffuse approximation model expressed in the frequency domain [1]:

$$-\nabla \cdot ([2(\kappa + \sigma)]^{-1} \nabla \varphi) + (\kappa + \frac{2\pi i \nu}{c}) \varphi = 0 \text{ in } \mathcal{D} \quad (7)$$

$$\varphi + \frac{A}{2\pi^{-1}} [2(\kappa + \sigma)]^{-1} \nabla \varphi \cdot \mathbf{n} = \frac{I}{\pi^{-1}} \mathbb{1}_{[\zeta \in \partial \mathcal{D}_s]} \text{ on } \partial \mathcal{D} \quad (8)$$

where  $\kappa$  and  $\sigma$  are the absorption and reduced scattering coefficients, respectively,  $c$  is the speed of light in the medium,  $\mathbf{n}$  is the unit outward normal vector,  $I$  is the total inward flux of the amplitude modulated diffuse source at the frequency  $\nu$ ,  $\mathbb{1}_{[\cdot]}$  denotes the indicator function,  $\partial \mathcal{D}_s$  depicts the light source location,  $i$  is the imaginary unit and  $A$  is a parameter which characterizes the reflection at the boundary and can be derived from Fresnel's law if specular reflection is considered [4] or from experimental set-ups [20].

The forward problem consists in solving eq. (7)-(8) for the photon density  $\varphi$  assuming the absorption coefficient  $\kappa$ , the reduced scattering coefficient  $\sigma$ , the location of the light source  $\partial \mathcal{D}_s$  and values of  $\nu$ ,  $c$ ,  $A$  and  $I$  are known. In contrast, the inverse problem consists in estimating the space-dependent radiative properties  $\kappa$  and  $\sigma$  on the basis of photon density measurements  $\check{\varphi}$  on  $\partial \mathcal{D}$ , under the assumption that the location of the light source  $\partial \mathcal{D}_s$  and values of  $\nu$ ,  $c$ ,  $A$  and  $I$  are known.

More specifically, the forward model leads to compute the state  $\varphi$  that depends on radiative properties  $\gamma = (\kappa, \sigma)$ . The difference (in the least squares sense) between this density and the measured one is integrated to the cost function to be minimized  $j(\gamma) = \mathcal{J}(\varphi)$ :

$$\mathcal{J}(\varphi) = \frac{1}{2} \sum_{k=1}^K \sum_{d=1}^D \left| \frac{\varphi(k;d) - \check{\varphi}(k;d)}{\check{\varphi}(k;d)} \right|^2 \quad (9)$$

where the index  $k$  defines the source (test) number and the index  $d$  defines the detection number. The solution of the inverse problem requires an optimization problem formulation of the kind: “Find the functions  $\kappa^*(\mathbf{x})$  and  $\sigma^*(\mathbf{x})$  such that  $j(\kappa^*, \sigma^*) = \min j(\kappa, \sigma)$ ”.

#### 3.2 Cost function derivative

In order to derive the Gauss–Newton (GN) algorithm, one first needs to recall basic definitions of directional derivatives. Following [17], let us denote the point  $(\kappa, \sigma) = \gamma \in \Lambda^2 \subset$

$[L_2(\mathcal{D})]^2$  and directions  $\eta$  and  $\zeta \in \Lambda$ . The directional derivative of the state  $\varphi$  at point  $\gamma$  towards  $\eta$  is

$$\varphi'(\gamma; \eta) := \lim_{\varepsilon \rightarrow 0^+} \frac{\varphi(\gamma + \varepsilon\eta) - \varphi(\gamma)}{\varepsilon} \quad (10)$$

An analogous definition can be stated for the directional derivative of the cost function along with its second-order directional derivative such that:

$$j''(\gamma; \eta, \zeta) := \lim_{\varepsilon \rightarrow 0^+} \frac{j'(\gamma + \varepsilon\zeta; \eta) - j'(\gamma; \eta)}{\varepsilon} \quad (11)$$

Next, the operators involved in Newton's method are extracted through the following equations:

$$j'(\gamma; \eta) = (\nabla j(\gamma), \eta)_{L_2(\mathcal{D})} \quad (12)$$

$$j''(\gamma; \eta, \zeta) = (\nabla^2 j(\gamma)\eta, \zeta)_{L_2(\mathcal{D})} \quad (13)$$

with the inner product is defined as  $(\gamma, \eta)_{L_2(\mathcal{D})} = \int_{\mathcal{D}} \gamma \eta \, d\mathbf{x}$ .

### 3.3 Parameterization

The solution of the optimization problem relies on iterative minimization algorithms for which control and state spaces must be finite in practice. To do so, let the functions to be determined  $\alpha = \kappa, \sigma$  and  $\varphi$  be approximated using finite element basis functions  $(\Theta_\xi)_{\xi=1}^{\Xi_\varphi}$  and  $(\psi_\xi)_{\xi=1}^{\Xi_\alpha}$ , respectively:

$$\varphi(\mathbf{x}) = \sum_{\xi=1}^{\Xi_\varphi} \Theta_\xi(\mathbf{x}) \tilde{\varphi}(\mathbf{x}_\xi) \quad (14)$$

$$\alpha(\mathbf{x}) = \alpha_{\text{ap}} \sum_{\xi=1}^{\Xi_\alpha} \psi_\xi(\mathbf{x}) \tilde{\alpha}(\mathbf{x}_\xi), \quad \alpha = \kappa, \sigma \quad (15)$$

where it should be noted that the dimension of the control space  $\Xi_\alpha$  is chosen less or equal to that of the state space  $\Xi_\varphi$ . As a consequence, projections of functions of the expression (15) into the state functional space must be used to make the solution of the system (7)-(8) possible. Moreover, parameters are adimensionalized with *a priori* properties  $\alpha_{\text{ap}} = \kappa_{\text{ap}}, \sigma_{\text{ap}}$  such that  $\alpha = \alpha_{\text{ap}} \tilde{\alpha}$ , in order to make both functions  $\tilde{\alpha} = \varkappa(\mathbf{x}), \varsigma(\mathbf{x})$  to be searched about unity.

Let us note  $\tilde{\gamma} = (\varkappa, \varsigma)$ . Parameterization of radiative property map functions allows the construction of the GN matrix system which is written as:

$$\tilde{\nabla}^2 j(\tilde{\gamma}) \delta \tilde{\gamma} = -\tilde{\nabla} j(\tilde{\gamma}) \quad (16)$$

with state second-order derivatives assumed to be negligible when compared to first-order state derivatives and where vector  $\tilde{\nabla} j$  and matrix  $\tilde{\nabla}^2 j$  represent continuous gradient  $\nabla j$  and

approached Hessian matrix  $\nabla^2 j$  of (12)-(13) decomposed on finite element basis functions  $(\psi_\xi)_{\xi=1}^{\Xi_\alpha}$ . Finally, the GN matrix system to be solved in order to update radiative properties at each iteration is expressed as:

$$\Re(\mathcal{S}^\top \bar{\mathcal{S}}) \delta\tilde{\gamma} = -\Re(\mathcal{S}^\top \bar{\mathcal{R}}) \quad (17)$$

where  $\mathcal{S}^\top$  and  $\bar{\mathcal{S}}$  are the transposed and the conjugate of  $\mathcal{S}$ , respectively, and  $\Re(\cdot)$  denotes the real part of the imaginary vector or matrix. The forward model behaving nonlinearly with respect to properties  $\tilde{\gamma} = (\varkappa, \varsigma)$ , the increment  $\delta\tilde{\gamma}^l = \tilde{\gamma}^l - \tilde{\gamma}^{l-1}$  given by the solution of the GN equation (17) is solved several times until convergence is attained. With the nomenclature defined earlier, the expressions for  $\mathcal{S}$  and  $\mathcal{R}$  are obtained such that:

$$\mathcal{S} = \left( \begin{array}{c} (\varphi'_\varkappa)_{(k;d)}^\xi, (\varphi'_\varsigma)_{(k;d)}^\xi \\ |\check{\varphi}_{(k;d)}|, |\check{\varphi}_{(k;d)}| \end{array} \right)_{\substack{\xi=1,\dots,\Xi_\alpha \\ (k;d)=(1,\dots,K;1,\dots,D)}} \quad (18)$$

$$\mathcal{R} = \left( \begin{array}{c} \varphi_{(k;d)} - \check{\varphi}_{(k;d)} \\ |\check{\varphi}_{(k;d)}| \end{array} \right)_{(k;d)=(1,\dots,K;1,\dots,D)} \quad (19)$$

where  $K$  and  $D$  are the number of sources and sensors, respectively.  $\mathcal{S} \in \mathbb{C}^{(K \times D) \times 2\Xi_\alpha}$ ,  $\mathcal{R} \in \mathbb{C}^{(K \times D)}$ , and derivative functions  $\varphi'_\varkappa$  and  $\varphi'_\varsigma$  can be obtained directly by differentiating the forward model (7)-(8) with respect to the adimensionalized radiative properties  $\varkappa$  and  $\varsigma$ , respectively.  $\varphi'_\varkappa$  and  $\varphi'_\varsigma$  are solution of the following partial differential equations:

$$-\nabla \cdot ([2(\kappa + \sigma)]^{-1} \nabla \varphi'_\varkappa) + (\kappa + \frac{2\pi i\nu}{c}) \varphi'_\varkappa = -\kappa_{\text{ap}} \varkappa' \varphi - \kappa_{\text{ap}} \nabla \cdot \left( \frac{\varkappa' \nabla \varphi}{2(\kappa + \sigma)^2} \right) \text{ in } \mathcal{D} \quad (20)$$

$$\varphi'_\varkappa + \frac{A}{2\pi^{-1}} [2(\kappa + \sigma)]^{-1} \nabla \varphi'_\varkappa \cdot \mathbf{n} = \kappa_{\text{ap}} \frac{A}{2\pi^{-1}} \frac{\varkappa' \nabla \varphi}{2(\kappa + \sigma)^2} \nabla \varphi \cdot \mathbf{n} \text{ on } \partial \mathcal{D} \quad (21)$$

and

$$-\nabla \cdot ([2(\kappa + \sigma)]^{-1} \nabla \varphi'_\varsigma) + (\kappa + \frac{2\pi i\nu}{c}) \varphi'_\varsigma = -\sigma_{\text{ap}} \nabla \cdot \left( \frac{\varsigma' \nabla \varphi}{2(\kappa + \sigma)^2} \right) \text{ in } \mathcal{D} \quad (22)$$

$$\varphi'_\varsigma + \frac{A}{2\pi^{-1}} [2(\kappa + \sigma)]^{-1} \nabla \varphi'_\varsigma \cdot \mathbf{n} = \sigma_{\text{ap}} \frac{A}{2\pi^{-1}} \frac{\varsigma' \nabla \varphi}{2(\kappa + \sigma)^2} \nabla \varphi \cdot \mathbf{n} \text{ on } \partial \mathcal{D} \quad (23)$$

where  $\varkappa'$ ,  $\varsigma'$  denote the perturbations. Algorithm 1 gives some explanations about the construction of vector  $\mathcal{R}$  and matrix  $\mathcal{S}$  involved in (17).

### 3.4 Tikhonov regularization

The Tikhonov regularization method consists in adding a penalization term  $\vartheta \mathcal{J}^+$  which is based on priors to the cost to be minimized (9):

$$\mathcal{J}^+(\varkappa, \varsigma) = \frac{1}{2} \sum_{\tilde{\alpha}=\varkappa, \varsigma} \int_{\mathcal{D}} (\tilde{\alpha} - \tilde{\alpha}_{\text{ap}})^2 \text{d}\mathbf{x} \quad (24)$$

---

**Algorithm 1:** Construction of vector  $\mathcal{R}$  and matrix  $\mathcal{S}$  involved in (17).

---

**Input:**  $\tilde{\gamma} = (\varkappa, \varsigma)$ ;

Compute  $\kappa = \kappa_{\text{ap}}\varkappa$ ,  $\sigma = \sigma_{\text{ap}}\varsigma$ ;

**for**  $k \leftarrow 1$  **to**  $K$  **do**

    Compute the state variable  $\varphi$  solving (7)-(8);

    Extract the complex values of the photon density  $\varphi$  at the nodes of the sensors and fill the corresponding part to the source  $k$  in vector  $\mathcal{R}$  (19);

**for**  $\xi \leftarrow 1$  **to**  $\Xi_\alpha$  **do**

        Set perturbations  $\varkappa'$  and  $\varsigma'$  to the finite element basis function  $\psi_\xi$ ;

        Compute  $\varphi'_{\varkappa}$  and  $\varphi'_{\varsigma}$  solving (20)-(21) and (22)-(23), respectively;

        Extract the complex values of  $\varphi'_{\varkappa}$  and  $\varphi'_{\varsigma}$  at the nodes of the sensors and fill the corresponding part to the source  $k$  and the perturbation  $\xi$  in matrix  $\mathcal{S}$  (18);

**return**  $\mathcal{R}$ ,  $\mathcal{S}$ ;

---

$\vartheta$  being the Tikhonov parameter, and  $\tilde{\alpha}_{\text{ap}}$  being equal to 1 as explained above. This has been successfully used in the optical tomography area e.g. by [23, 5, 21] when using the GN optimizer. Doing so, the optimization problem becomes “*Find the functions  $\varkappa_\vartheta^*(\mathbf{x})$  and  $\varsigma_\vartheta^*(\mathbf{x})$  such that  $j_\vartheta(\varkappa_\vartheta^*, \varsigma_\vartheta^*) = \min j_\vartheta(\varkappa, \varsigma)$ ” with  $j_\vartheta = \mathcal{J} + \vartheta\mathcal{J}^+$ .*

The simultaneous use of several regularizations together has been studied in [7, 2], among which the reduction of the control space dimension [3] with Tikhonov regularization. It is shown here that the control space reduction modifies only slightly the optimal Tikhonov parameter.

The control space mesh is assumed to be chosen sufficiently fine to ensure that the spatial fluctuations of the radiative properties are well taken into account. Thus, the projections from the control functional space to the state functional space are such that the state is almost independent of the number of degrees of freedom  $\Xi_\alpha$ , thus the cost function  $\mathcal{J}$  is also almost independent of  $\Xi_\alpha$ . Next, to simplify the presentation, let us consider only one parameter, say  $\varkappa$ .

**Proposition 3.1.** *Under the following assumptions: H1)  $\psi_\xi$  are basis functions constant per element on a regular mesh, i.e. Lagrange  $\mathcal{P}_0$  elements are used for the discretization of the function  $\varkappa$ , H2)  $\varkappa \sim \mathcal{N}(1, \sigma_\varkappa^2)$ , then the optimal Tikhonov parameter is asymptotically independent of the dimension  $\Xi_\alpha$ .*

*Proof.* Let  $\varkappa_\vartheta^*$  be a solution of the optimization problem

$$\varkappa_\vartheta^* = \arg \min_{\varkappa} (\mathcal{J}(\varphi) + \vartheta\mathcal{J}^+(\varkappa)) \quad (25)$$

with  $\vartheta$  representing the best compromise between both costs  $\mathcal{J}(\varphi)$  and  $\mathcal{J}^+(\varkappa)$ . Then, showing  $\mathcal{J}^+$  does not depend on the parameterization yields  $\vartheta$  independent of  $\Xi_\alpha$ . Let  $\tilde{\vartheta} = \sigma_\varkappa^2\vartheta$  and  $\Delta x = |\mathcal{D}|/\Xi_\alpha$  be the area of one element of the regular mesh. With H1), we have

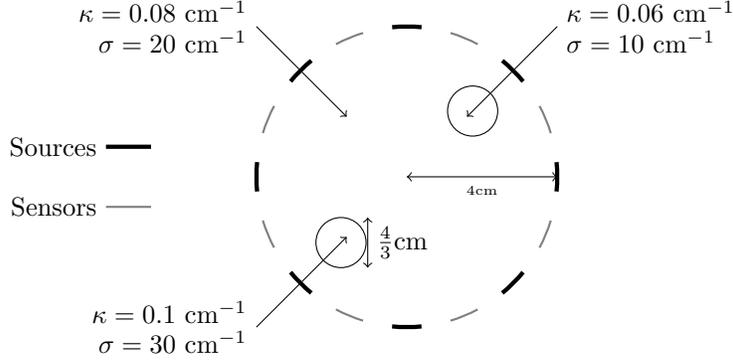


Figure 3: Test medium geometry representation. Eight sources are located on the boundary. For each source (which constitutes a test), the emerging radiation is measured on all sensors.

$\vartheta \mathcal{J}^+(\varkappa) = \frac{\vartheta |\mathcal{D}|}{2} \frac{1}{\Xi_\alpha} \sum_{i=1}^{\Xi_\alpha} \left( \frac{\varkappa_i - 1}{\sigma_\varkappa} \right)^2$ . With H2),  $\frac{\varkappa_i - 1}{\sigma_\varkappa} \sim \mathcal{N}(0, 1) \forall i$ , thus  $\Upsilon = \sum_{i=1}^{\Xi_\alpha} \left( \frac{\varkappa_i - 1}{\sigma_\varkappa} \right)^2 \sim \chi_{\Xi_\alpha}^2$ . Then, using the fact that  $\mathbf{E}(\Upsilon) = \Xi_\alpha$ , we have  $\mathbf{E}(\mathcal{J}^+) = \frac{\sigma_\varkappa^2 |\mathcal{D}|}{2}$  which is independent of  $\Xi_\alpha$ . This completes the proof.  $\square$

*Remark 3.2.* This demonstration can be easily extended to several parameters, say with a function of the type of (24) assuming that  $\sigma_\varsigma \approx \sigma_\varkappa$ .

*Remark 3.3.* It could also be extended to other finite element parameterizations, but with much heavier calculations.

### 3.5 Numerical validation

The ability to determine on a coarse mesh a reasonable parameter for the Tikhonov regularization is illustrated on a circular test medium. The rationale behind this choice is to reduce computational time required for getting a Tikhonov parameter so that the solution is close enough to the true solution, at least from an engineering point of view.

Synthetic data and predicted values are collected with the help of 8 equally distributed sources and sensors about the disk perimeter ( $K = 8$ ). Sources and sensors are alternatively distributed as schematically depicted in Figure 3. Each sensor contains 7 pointwise photon density measurements ( $D = 56$ ). Radiative property maps to be recovered are also given in Figure 3. The areas to be detected are located at about 2.5 cm in opposite locations at 45° and 225°, respectively. The defects are 4/3 cm in diameters while the whole medium has a 4 cm radius. Each probe is 0.8 cm in length. The other physical parameters involved in (7)-(8) are fixed to:  $n = 1.4$ ,  $\nu = 1 \times 10^8 \text{ s}^{-1}$ ,  $c = c_0/n$  with  $c_0 = 3 \times 10^8 \text{ m.s}^{-1}$ ,  $I = 1 \text{ W.m}^{-2}$  and  $A$  is derived from Fresnel's law (see [4]).

Radiative parameters are initialized to match the medium background considered as priors:  $\kappa^0 = \kappa_{\text{ap}} = 0.8 \text{ cm}^{-1}$  and  $\sigma^0 = \sigma_{\text{ap}} = 20 \text{ cm}^{-1}$  all over the medium  $\mathcal{D}$ , implying  $\varkappa = \varsigma = 1$ . Synthetic data  $\check{\varphi}$  and initial adimensionalized radiative properties  $\varkappa$  and  $\varsigma$

	$\Xi_\alpha = \dim \mathcal{V}_h^\alpha$	$\dim \mathcal{V}_h^\varphi$	$\dim \mathcal{V}_h^{\check{\varphi}}$
case 1	2924	2924	5336
case 2	1382		
case 3	384		

Table 2: Dimensions of finite element spaces for  $\varphi$ ,  $\check{\varphi}$  and  $\alpha$ .

are inputs of the inverse problem. The way in which synthetic data are generated is given hereafter.

Within the inverse procedure, the forward model (7)-(8) is solved by a finite element method with Lagrange  $\mathcal{P}_1$  elements, that is with  $\varphi \in \mathcal{V}_h^\varphi = \mathcal{V}(\mathcal{M}_h^\varphi, \mathcal{P}_1)$ . The state mesh  $\mathcal{M}_h^\varphi$  is chosen fine enough to ensure that numerical predictions can fit the synthetic data. Synthetic data  $\check{\varphi}$  are also generated with the finite element method with Lagrange  $\mathcal{P}_1$  elements but using a much finer mesh  $\mathcal{M}_h^{\check{\varphi}}$  in order to avoid the inverse crime [16]. After projecting the synthetic data in the functional space  $\mathcal{V}_h^\varphi$ ,  $\check{\varphi}$  are corrupted with a multiplicative white Gaussian noise of signal-to-noise ratio SNR = 30 dB:  $\check{\varphi}_{\text{noisy}} = \check{\varphi} (1 + 10^{-\text{SNR}/10} \times \varepsilon)$  where  $\varepsilon \sim \mathcal{N}(0, 1)$ .

Next, the control space  $\mathcal{V}_h^\alpha = \mathcal{V}(\mathcal{M}_h^\alpha, \mathcal{P}_1)$  for  $\alpha = \kappa, \sigma$  relies on a more or less coarse mesh  $\mathcal{M}_h^\alpha$  according to cases (see Table 2). The FreeFem++ environment has been used to perform all computations [13].

The distance curves  $E(\vartheta)$  are built for the three cases described in Table 2: a fine mesh for the state ( $\dim \mathcal{V}_h^\varphi = 2924$ ), and successively a fine mesh, a medium mesh and a coarse mesh with respectively  $\dim \mathcal{V}_h^\alpha$  equal to 2924, 1382 and 384 nodes for the parameters. The three distance curves are presented all together in Figure 4. Figure 4 first shows that for  $\vartheta > 0.4$ , the calculated errors are independent of the mesh coarseness. Otherwise, for  $\vartheta < 0.3$ , it is seen that the reduction of the control space reduces dramatically the errors. Such reduction of control space thus regularizes the inverse problem preconditioning the optimization [3]. In between, Figure 4 shows that the optimal Tikhonov parameters are found to be quasi-independent of the number of degrees of freedom of the finite control space. Within the range  $[10^{-2}, 10^{+2}]$ , the optimal Tikhonov parameter  $\vartheta^*$  is found to equal 0.3 for  $\Xi_\alpha = 384$  and 0.4 for  $\Xi_\alpha = 1382$  and 2924. This means that the minimal residual is indeed quasi-independent of the mesh coarsening, i.e. of  $\dim \mathcal{V}_h^\alpha$ . This is what has been shown in Proposition 3.1.

Table 3 also provides the relative CPU time needed to build such curves for cases 1, 2 and 3 described in Table 2: the construction of this error curve with the coarsest mesh needs, in this case, roughly a tenth of the initial CPU time related to the construction with the finest mesh. This can be easily understood by examining the construction scheme of the GN matrix system presented above (see Algorithm 1).

Figures 5 and 6 present the reconstructions of both  $\kappa$  and  $\sigma$  with the GN optimizer, at the 20<sup>th</sup> iteration, and with the Tikhonov weight equal to 0.2 and 0.4, respectively. Figure 5 shows the divergence of the reconstruction with negative values for  $\sigma$  due to under-regularization: the weight parameter  $\vartheta$  is too low. Figure 6 presents the same reconstruction but with the Tikhonov parameter  $\vartheta$  chosen at the minimum of the error curve. Reconstruc-

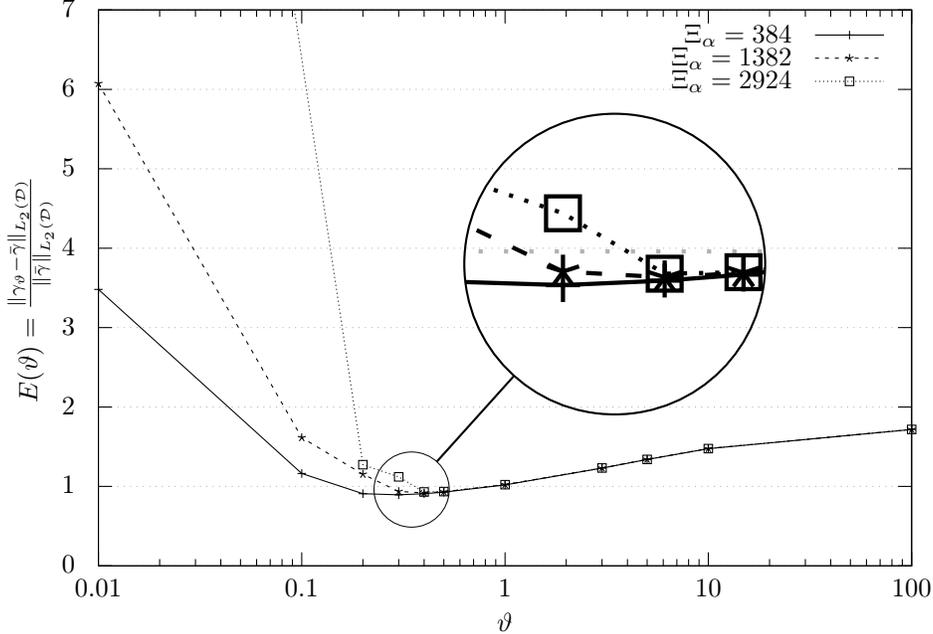


Figure 4: Distance from the solution to the target  $E(\vartheta)$  as a function of the Tikhonov parameter  $\vartheta$ , for three control space dimensions  $\Xi_\alpha$  equal to 384, 1382 and 2924. It is seen that  $\vartheta^* = \arg \min E(\vartheta)$  is quasi-independent of  $\Xi$ .

tions are good even in the presence of a 30 dB noise.

	Case 1	Case 2	Case 3
CPU time	1	0.38	0.09

Table 3: CPU time comparisons of L-Curve generation

## 4 Conclusion

This paper has demonstrated that the optimal Tikhonov parameter, defined as that which minimizes the distance between the actual solution and the noise-free solution, is almost independent of the control space parameterization. This has first been verified on a standard linear inverse heat conduction problem in which, in phase one, the Tikhonov parameter has been approximated with the singular value decomposition and the maximum discrepancy principle and then, in a second phase, in the sense of the optimal Tikhonov parameter previously defined. Next, this validation has been extended to the non-linear inverse problem of diffuse optical tomography. The error curves with respect to the Tikhonov parameter showed that their minimum are quasi-independent of the control space parameterization. This may result in CPU time reduction when searching for the optimal Tikhonov parameter,

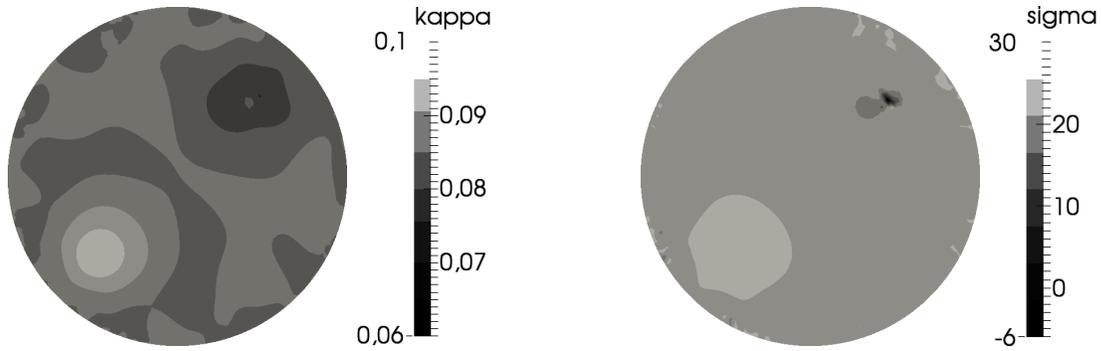


Figure 5: Reconstruction of  $\kappa$  (left) and  $\sigma$  (right) with  $\vartheta = 0.2$ . It should be noted the divergence for the reduced diffusion coefficient.

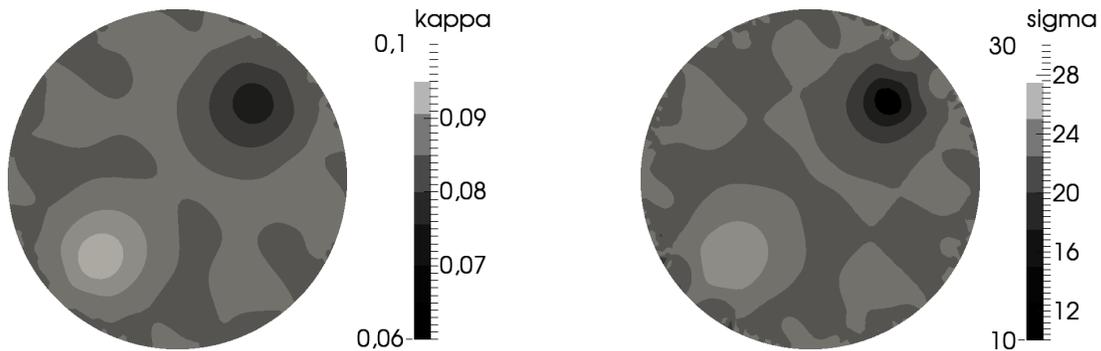


Figure 6: Reconstruction of  $\kappa$  (left) and  $\sigma$  (right) with  $\vartheta = 0.4$

whatever the method used to find it. Of course, this is highly recommended for large size (three-dimensional) objects when the Gauss–Newton optimizer is chosen. Finally, note that the two inverse problems dealt with in this paper were under-determined, as it is the case in most space-dependent function estimation problems. Consequently, the theory has been numerically validated only on these two specific inverse problems. This may be viewed as a limitation, but the need of control space dimension reduction is much less useful for over-determined inverse problems than for space-dependent ones which are usually under-determined.

## 5 Acknowledgements

The authors would like to thank Pr. Hecht for providing the very efficient finite element tool FreeFem++ [13]. The authors would also like to thank the Levis’s town, the Regional Conference of Elected Representatives of Chaudière-Appalaches, Énergie Valero Inc. and Ecosystem for their financial support to the industrial research chair t3e. The authors also thank the Centre Régional de Calcul Intensif des Pays de la Loire (CCIPL), financed by the French Research Ministry, the Région Pays de la Loire and the University of Nantes, allowing computations on their supercomputers. Finally, the authors would like to sincerely thank

the reviewers for their constructive and relevant comments formulated on the first version that significantly enhanced the quality of this paper.

## References

- [1] Simon R Arridge. Optical tomography in medical imaging. *Inverse problems*, 15(2):R41, 1999.
- [2] O Balima, Y Favennec, and D Rouse. Optical tomography reconstruction algorithm with the finite element method: An optimal approach with regularization tools. *Journal of Computational Physics*, 251:461–479, 2013.
- [3] Guy Chavent. *Nonlinear least squares for inverse problems: theoretical foundations and step-by-step guide for applications*. Springer, 2010.
- [4] H. Dehghani, S. Srinivasan, B. W. Pogue, and A. Gibson. Numerical modelling and image reconstruction in diffuse optical tomography. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1900):3073–3093, 2009.
- [5] Hamid Dehghani, Matthew E Eames, Phaneendra K Yalavarthy, Scott C Davis, Subhadra Srinivasan, Colin M Carpenter, Brian W Pogue, and Keith D Paulsen. Near infrared optical tomography using nifast: Algorithm for numerical model and image reconstruction. *Communications in numerical methods in engineering*, 25(6):711–732, 2009.
- [6] GS Dulikravich and TJ Martin. Inverse shape and boundary condition problems and optimization in heat conduction. *Advances in Numerical Heat Transfer*, 1:381–426, 1996.
- [7] Yann Favennec, Fabien Dubot, Benoit Rousseau, and Daniel Rouse. Mixing regularization tools for enhancing regularity in optical tomography applications. In *IPDO-2007-Inverse Problems, Design and Optimization Symposium*, 2013.
- [8] Gene H Golub, Michael Heath, and Grace Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21(2):215–223, 1979.
- [9] AV Goncharskii, Aleksandr Sergeevich Leonov, and Anatolii Grigor’evich Yagola. A generalized discrepancy principle. *USSR Computational Mathematics and Mathematical Physics*, 13(2):25–37, 1973.
- [10] Xuejun Gu, Yong Xu, and Huabei Jiang. Mesh-based enhancement schemes in diffuse optical tomography. *Medical physics*, 30(5):861–869, 2003.
- [11] Jacques Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton university bulletin*, 13(49-52):28, 1902.

- [12] Per Christian Hansen and Dianne Prost O’Leary. The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM Journal on Scientific Computing*, 14(6):1487–1503, 1993.
- [13] F Hecht. New development in freefem++. *Journal of Numerical Mathematics*, 20(3-4):251–266, 2012.
- [14] Edward Hensel and Richard Hills. Steady-state two-dimensional inverse heat conduction. *Numerical Heat Transfer, Part B Fundamentals*, 15(2):227–240, 1989.
- [15] Y. Jarny. Lecture 9: Inverse problems & regularized solutions. Eurotherm Spring School METTI 2011 : Thermal measurements and inverse techniques: <http://www.sft.asso.fr/document.php?pagendx=12299&project=sft>, Roscoff, June 2011.
- [16] Jari Kaipio and Erkki Somersalo. *Statistical and computational inverse problems*, volume 160. Springer, 2005.
- [17] Jacques-Louis Lions and Pierre Faurre. *Cours d’analyse numérique*. École Polytechnique, 1982.
- [18] Vladimir Alekseevich Morozov, Z Nashed, and AB Aries. *Methods for solving incorrectly posed problems*. Springer, 1984.
- [19] Dianne P O’Leary. Near-optimal parameters for tikhonov and other regularization methods. *SIAM Journal on scientific computing*, 23(4):1161–1171, 2001.
- [20] M Schweiger, SR Arridge, M Hiraoka, and DT Delpy. The finite element method for the propagation of light in scattering media: boundary and source conditions. *Medical physics*, 22(11):1779–1792, 1995.
- [21] Martin Schweiger, Simon R Arridge, and Ilkka Nissilä. Gauss–newton method for image reconstruction in diffuse optical tomography. *Physics in medicine and biology*, 50(10):2365, 2005.
- [22] J Taler. Nonlinear steady-state inverse heat conduction problem with space-variable boundary conditions. *Journal of heat transfer*, 114(4):1048–1051, 1992.
- [23] T Tarvainen, M Vauhkonen, and SR 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):2767–2778, 2008.
- [24] AN Tikhonov, AS Leonov, and AG Yagola. *NonLinear Ill-Posed Problems, Vol. 1 and 2, 1998*. London, Chapman and Hall, 1998.
- [25] Curtis R Vogel. Non-convergence of the l-curve regularization parameter selection method. *Inverse problems*, 12(4):535, 1996.

- [26] AG Yagola, AS Leonov, and VN Titarenko. Data errors and an error estimation for ill-posed problems. *Inverse problems in engineering*, 10(2):117–129, 2002.