Regularization Techniques for Inverse Problem in DOT Applications

To cite this article: A Benfenati *et al* 2020 *J. Phys.: Conf. Ser.* **1476** 012007

View the article online for updates and enhancements.

Regularization Techniques for Inverse Problem in DOT Applications

A Benfenati[1], P Causin[2], MG Lupieri[3], G Naldi[4]

[1,4] Dipartimento di Scienze Politiche ed Ambientali, Universitá degli studi di Milano, Via Celoria 2, 20133, Milano
[2,3] Dipartimento di Matematica "Federigo Enriques", Universitá degli studi di Milano, Via C. Saldini 50, 20133, Milano

E-mail: [1]`alessandro.benfenati@unimi.it`, [2]`paola.causin@unimi.it`, [3]`marinagiulia.lupieri@unimi.it`, [4]`giovanni.naldi@unimi.it`

Abstract. Diffuse optical tomography (DOT) is an emerging diagnostic technique which uses near–infra–red light to investigate the optical coefficients distribution in biological tissues. The surface of the tissue is illuminated by light sources, then the outgoing light is measured by detectors placed at various locations on the surface itself. In order to reconstruct the optical coefficients, a mathematical model of light propagation is employed: such model leads to the minimization of the discrepancy between the detected data and the corresponding theoretical field. Due to severe ill–conditioning, regularization techniques are required: common procedures consider mainly ℓ_1–norm (LASSO) and ℓ_2–norm (Tikhonov) regularization. In the present work we investigate two original approaches in this context: the *elastic–net* regularization, previously used in machine learning problems, and the Bregman procedure. Numerical experiments are performed on synthetic 2D geometries and data, to evaluate the performance of these approaches. The results show that these techniques are indeed suitable choices for practical applications, where DOT is used as a cheap, first–level and almost real–time screening technique for breast cancer detection.

1. Introduction

Diffuse optimal tomography (DOT) is a technique that received a growing interest in recent years. It consists in a non–invasive and non–ionizing imaging procedure that enables to investigate absorption and scattering optical properties of biological tissue. It uses near–infra–red light (600–900 nm): in this particular spectral window the absorption properties of water, oxygenated hemoglobin (HbO) and de–oxygenated hemoglobin (Hb) allow the light to penetrate deeper to examine the tissue up to several centimetres in depth [10, 24]. The concentration of HbO and Hb produce a significant difference in optical coefficients between healthy and diseased tissue: such difference provides an imaging contrast tool for tissue diagnostic.

The technical arrangement of DOT consists in several detectors recording the light intensity from differently located sources. The light emerging from the sample is collected by an array of photodetectors or a CCD camera, depending on the specific industrial technology (see [10] for more details). The focus of this work is on the inverse problem in DOT, assuming that the image pre-processing and filtering has already been performed by the recording hardware and/or specialized software. The aim is reconstructing the distribution of the tissue optical coefficient. This optical distribution reveals to be the solution of a minimization problem,

where the objective functional measures the discrepancy between the observed data and the corresponding optical field given by a light propagation model. The ill–conditioning of the problem, due to scattering properties of the tissue, requires the use of regularization techniques [1, 2] in order to obtain a reliable distribution.

Natural choices for the regularization functional are ℓ_2 norm and ℓ_1 norm. The former, known also as Tikhonov regularization or ridge regression, has been widely used in DOT [25, 12, 18]. The ℓ_1–norm [27, 15, 9] promotes sparse solutions: it forces some components on the solution to strongly emerge and at the same time pushes others to very small values or even to zero. In DOT application the ℓ_1 norm has been used to improve sparsity and sharp edges detection [26]. On one hand, the ℓ_1–norm is indeed a suitable choice, since it allows to emphasize the contrast between the homogeneous and healthy tissue and pathological formations, usually small and localized. On the other hand, this norm selects only a small number of variables from a group of correlated ones. Moreover, the presence of noise on the recorded data amplifies these aspects: thus, the use of the sole ℓ_1–norm could lead to imprecisions on localizing the inclusions and on reconstructing their shape.

In this work two novel regularization approaches are proposed. The first one is based on the elastic–net regularization [29], which consist in a linear combination of ℓ_1 and ℓ_2 norms and it was firstly proposed in the context of machine learning problems. This functional selects groups of correlated variables (ℓ_2 term), yielding thus a good localization, and produces a realistic sparse pattern in the solution (ℓ_1 term). Whilst this technique has been largely used in machine learning [13], at the best of our knowledge, use in DOT applications is original [11]: beside the already mentioned ℓ_2 [28] and ℓ_1 [23] norms, the interested reader can find in the literature other several approaches, such as algebraic (ART, SIRT) and subspaces (TSVD) techniques, TCG methods [17], but the combination of Tikhonov and ℓ_1 is a novelty. The second approach in this context is the usage of the inexact Bregman procedure [4, 6, 30]: such procedure consists in substituting the regularization functional with its Bregman distance, triggering an iterative procedure which usually yields to contrast enhancement [7] and sometimes a minor computational time [4]. Both aspects are crucial: the former satisfy the objective of improving the visual inspection of the different absorption coefficients, while the latter contributes in creating an almost real–time screening technique. Some authors [21] had applied the Linearized Bregman technique, but actually this approach amounts to solve a different problem than the original [30].

The paper is organized as follows. In Section 2 DOT forward and backward problems are presented. Section 3 is dedicated to the regularization approaches. Section 4 is devoted to the numerical experimentation, where the performance of the proposed techniques is evaluated. Section 5 draws the conclusions of the present work.

2. Mathematical models for light propagation

2.1. Forward problem

An exhaustive but computationally intensive mathematical model of light propagation in turbid media is represented by the Radiative Transfer Equation (RTE), an integro–differential equation which ensures the conservation of energy of the light radiance ($[\text{W/cm}^2\text{sr}]$). Performing an expansion in spherical harmonics of the radiance provides a series of approximating equations: its truncation to the first N terms yields the so–called P_N approximation.

Let $\mathbf{r} \in \Omega \subset \mathbb{R}^3$ be the position in the domain Ω (linear unit in [cm]) and $S(\mathbf{r})$ $[\text{W/cm}^3]$ the isotropic volume source. Considering a P_1 expansion of the RTE and under the additional hypotheses of isotropic sources and slow variation of the photonic flux, the following diffusive equation (DE) is obtained for the fluence rate at steady–state (see [14] for a complete derivation).

$$-\nabla \cdot (D(\mathbf{r})\nabla\Phi(\mathbf{r})) + \mu_a(\mathbf{r})\Phi(\mathbf{r}) = S(\mathbf{r}) \tag{1}$$

where the photon fluence Φ $[\text{W/cm}^2]$ represents the integral of the radiance over the solid angle

in three dimensional space. Moreover, D [cm] is the diffusion coefficient defined as:

$$D(\mathbf{r}) = \frac{1}{3(\mu_a(\mathbf{r}) + \mu_s'(\mathbf{r}))} \approx \frac{1}{3\mu_s'(\mathbf{r})}$$

where $\mu_s'(\mathbf{r}) = \mu_s(\mathbf{r})(1 - g)$ [1/cm] is the reduced scattering coefficient, μ_s being the scattering coefficient, g the anisotropy scattering factor and $\mu_a = \mu_a(\mathbf{r})$ [1/cm] is the absorption coefficient, with $\mu_s \gg \mu_a$. From the microscopic viewpoint, in the DE modeling approach, the photons are supposed to move throughout the sample along random paths. Each photon travels along straight segments with a sudden discontinuity when the photon changes direction or is absorbed. The average length of rectilinear tracts l_{tr} is the mean free path, $l_{tr} \approx \frac{1}{\mu_s}$.

2.2. Inverse problem
We consider the reduced scattering coefficient μ_s' to be the same inside the target, so that $D \approx D_0 = \text{const}$. In this case, the goal of DOT is to reconstruct $\mu_a(\mathbf{r})$ from a set of boundary measurements of $\Phi(\mathbf{r})$, assuming D is known. In the following, the Rytov perturbative approach is adopted. This choice is acceptable in the hope of disposing of fast algorithms to provide "almost real time" diagnostic results. Namely

$$\Phi(\mathbf{r}) = e^{\psi(\mathbf{r})} = e^{\psi_0(\mathbf{r}) + \psi_1(\mathbf{r})} = \Phi_0 e^{\psi_1(\mathbf{r})} \tag{2}$$

where a series solution for $\psi(\mathbf{r})$ can be derived, ψ_0 (and Φ_0) being the so–called background conditions (in absence of perturbations) and ψ_1 the perturbed field induced by the heterogeneities. After some algebra (see [20], Vol.II, Ch.17 for a detailed derivation), the following equation for $(\Phi_0\psi_1)$ is obtained

$$\left[\Delta - \frac{\mu_a}{D_0}\right](\Phi_0\psi_1) = \left(\frac{\delta\mu_a}{D_0} - \nabla\psi_1 \cdot \nabla\psi_1\right)\Phi_0, \tag{3}$$

where $\mu_a(\mathbf{r}) = \mu_{a,0} + \delta\mu_a(\mathbf{r})$, $\delta\mu_a$ being a perturbation from background conditions. A linearized solution $\psi_1 \sim \psi_{1,0}$ for problem (3) is obtained from

$$\psi_{1,0}(\mathbf{r}) = \frac{1}{\Phi_0(\mathbf{r})} \int_\Omega G(\mathbf{r} - \mathbf{r}')\frac{\delta\mu_a(\mathbf{r}')}{D_0}\Phi_0(\mathbf{r}')\,d\mathbf{r}', \tag{4}$$

$G(\mathbf{r} - \mathbf{r}')$ being the Green's function of (3) evaluated in $\mathbf{r}$ for a Dirac–δ source located in $\mathbf{r}'$.

Notice that $\psi_1 = \log(\Phi/\Phi_0)$ represents the *logarithmic amplitude fluctuation* of the light intensity and it is the datum used as a rhs in the inverse problem solution.

The integral relation (4) provides an explicit model solution and is usually dealt with by discretizing the domain into a large number N of volume elements (voxels) and evaluating the rhs of the equation at the detector positions. Using for example a midpoint quadrature rule (voxel centroid $\mathbf{r}_j$, voxel volume ΔV_j, $j = 1,\ldots,N$) to discretize Eq. (4), one obtains

$$\psi_{1,0}(\mathbf{r}^s, \mathbf{r}^d) \simeq \sum_{j=1}^{N} \frac{\Delta V_j}{\Phi_0(\mathbf{r}^s, \mathbf{r}^d)}G(\mathbf{r}^d - \mathbf{r}_j)\frac{\delta\mu_a(\mathbf{r}_j)}{D_0}\Phi_0(\mathbf{r}^s, \mathbf{r}_j), \tag{5}$$

where the dependence on the source and detector position $\mathbf{r}^s$ and $\mathbf{r}^d$, respectively, has been made explicit. Considering now to dispose of $M = N_s \times N_d$ measurements (that is, M couples source–detector, N_s is the number of the sources, N_d the number of detectors) the following system of equations naturally arises

$$\sum_{j=1}^{N} \frac{\Delta V_j}{\Phi_0(\mathbf{r}_i^s, \mathbf{r}_i^d)}G(\mathbf{r}_i^d - \mathbf{r}_j)\frac{\delta\mu_a(\mathbf{r}_j)}{D_0}\Phi_0(\mathbf{r}_i^s, \mathbf{r}_j) = \log\left(\frac{\Phi(\mathbf{r}_i^s, \mathbf{r}_i^d)}{\Phi_0(\mathbf{r}_i^s, \mathbf{r}_i^d)}\right) \qquad i = 1,\ldots,M, \tag{6}$$

which can be put in matricial form as

$$\mathbf{J}\,\boldsymbol{\delta\mu_a} = \mathbf{y}, \tag{7}$$

where $J_{ij} = \dfrac{\Delta V_j}{\Phi_0(\mathbf{r}_i^s, \mathbf{r}_i^d)} G(\mathbf{r}_i^d - \mathbf{r}_j)\dfrac{1}{D_0}\Phi_0(\mathbf{r}_i^s, \mathbf{r}_j) \in \mathbb{R}^{(M \times N)}$ is the Jacobian (*sensitivity matrix*),

$\boldsymbol{\delta\mu_a} \in \mathbb{R}^{(N \times 1)}$ is the unknown vector and $\mathbf{y}_i = \log\left(\dfrac{\Phi(\mathbf{r}_i^s, \mathbf{r}_i^d)}{\Phi_0(\mathbf{r}_i^s, \mathbf{r}_i^d)}\right) \in \mathbb{R}^{(M \times 1)}$ is obtained from

measurements or from numerical simulations of the forward model.

The inversion of system (7) is a nontrivial task, since the system can be overdetermined ($M > N$) or underdetermined ($M < N$) in the case of strong data deficiency and the matrix $\mathbf{J}$ is ill conditioned. In the forthcoming sections, we discuss the issue of regularization procedures applied to the solution of (7). To do this, it is important to observe that problem (7) stems from the minimization procedure of the Least Square functional

$$\min_{\boldsymbol{\delta\mu_a}} \frac{1}{2}\|\mathbf{y} - \mathbf{J}\boldsymbol{\delta\mu_a}\|_2^2. \tag{8}$$

3. Regularization Techinques

3.1. Elastic Net Penalization

The functional in Eq. (8) is regularized via the *elastic–net* [29]

$$\min_{\boldsymbol{\delta\mu_a}} \frac{1}{2M}\|\mathbf{y} - \mathbf{J}\boldsymbol{\delta\mu_a}\|_2^2 + \lambda\left(\alpha\|\boldsymbol{\delta\mu_a}\|_1 + (1-\alpha)\|\boldsymbol{\delta\mu_a}\|_2^2\right). \tag{9}$$

The parameter λ is a regularization parameter and α controls the influence of the penalties given by the ℓ_1 and ℓ_2 norms on the computed solution. More precisely, λ balances the trade–off between the Least Square functional and the regularization, while α controls the convex combination of the two norms, allowing to decide which one has the greater influence ($\alpha = 0$ sets the Thikonov regularization, $\alpha = 1$ leads to ℓ_1 regularization). By varying $\alpha \in (0, 1)$ sparsity is traded for the inclusion of more solution components contributions. The choice of λ is quite delicate. For pure sparsity–promotion via ℓ_1, there is still not any golden rule for an automatic selection for λ: some specific strategies refer to particular instances [3, 22]. For the pure ℓ_2 penalization there exist several more consolidated methods for automatic choice of λ [25]. A general possibility, implemented in advanced software packages as the `glmnet` [16] package, is to use k−fold cross validation techniques to establish optimal values of λ.

3.2. Bregman Procedure.

The Bregman procedure [4] finds its basis on the Bregman distance: given a convex function f with a non–empty subdifferential ∂f, its Bregman distance from $\mathbf{x}$ to $\mathbf{y}$ is $D_f^p(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) - f(\mathbf{y}) - \langle \mathbf{p}, \mathbf{x} - \mathbf{y}\rangle$, $\mathbf{p} \in \partial f(\mathbf{y})$, being $\partial f(\mathbf{y})$ the subdifferential of f at $\mathbf{y}$ [8, 9]. When $f \in \mathcal{C}^1$, $\partial f(\mathbf{y}) \equiv \nabla f(\mathbf{y})$. In this work, the Least Square functional is coupled with ℓ_2 and ℓ_1 regularization, triggering the following scheme for the Tikhonov regularization

$$\boldsymbol{\delta\mu_a}^{k+1} = \left(\mathbf{J}^\top \mathbf{J} + \lambda \boldsymbol{I_d}\right)^{-1}\left(\mathbf{J}^\top \mathbf{y} + \lambda\boldsymbol{\delta\mu_a}^k\right), \quad \boldsymbol{\delta\mu_a}^0 = 0 \tag{10}$$

while for LASSO regularization the algorithm reads as

$$\begin{cases} \boldsymbol{\omega}^{k,0} = \boldsymbol{\delta\mu_a}^k \\ \begin{cases} \boldsymbol{\omega}^{k,\ell+1/2} = \left(\boldsymbol{I_d} - \mathbf{J}^\top \mathbf{J}\right)\boldsymbol{\omega}^{k,\ell} + \gamma\mathbf{J}^\top \mathbf{y} + \gamma\lambda\mathbf{p}^k \\ \boldsymbol{\omega}^{k,\ell+1} = \text{soft}_{\gamma\lambda}\left(\boldsymbol{\omega}^{k,\ell+1/2}\right) \end{cases} \\ \boldsymbol{\delta\mu_a}^{k+1} = \boldsymbol{\omega}^{k,\ell+1} \\ \mathbf{p}^{k+1} = \mathbf{p}^k - \frac{1}{\lambda}\mathbf{J}^\top\left(\mathbf{J}\boldsymbol{\delta\mu_a}^{k+1} - y\right) \end{cases} , \quad \boldsymbol{\delta\mu_a}^0 = 0,\ \mathbf{p}^0 = 0,\ \gamma > 0 \tag{11}$$

where $\text{soft}_\gamma(x) = \text{sign}(x) \max(|x| - \gamma, 0)$. In [4, 6] the convergence proof of the general method is stated and more details are given in [30] for the particular case of Eq. (11).

4. Numerical Tests

A 2D domain Ω is considered: it consists in a semicircle of radius 5 cm (Fig. 1). This geometry, albeit in a reduced dimensional setting and idealized version, represents a female breast supported on a solid plate on which embedded light sources are alternatively turned on for the DOT exam.
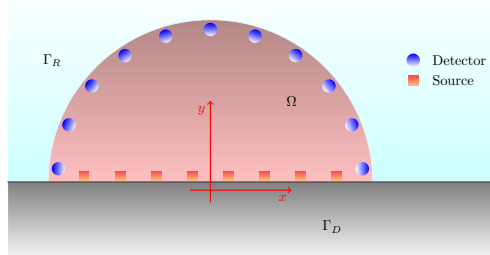

Figure 1. Common geometry and source/detector positioning for the numerical experiments. The scheme depicts 2D case, the 3D one follows straightforwardly the same idea. Γ_D represents the surface on which the tissue is leaning, Γ_R the open air surrounding the surface Ω, $\partial\Omega = \Gamma_D \cup \Gamma_R$.

The synthetic fluence observations in the detectors are generated by solving (1) with the MatLab® PDE Toolbox on a fine triangular mesh consisting of almost 37858 triangles and using an approximation of degree 2. The boundary conditions are $\varphi = 0$ on $\partial\Gamma_D$, $\varphi + 2\,A\,D\,\partial\varphi/\partial\vec{n} = 0$ on Γ_R, $\vec{n}$ is the outward normal vector and the coefficient A keeps into account the different tissue and air refractive indices [14]. The first condition represents the fact that light does not escape from the solid plate. $N_s = 20$ point–wise light sources are simulated, uniformly located on the horizontal boundary at a distance of 1 mm from Γ_D, $N_d = 200$ detectors are radially disposed 1 mm inside the domain along the boundary Γ_R. We discretize the integral using N = 3882 voxels. The background values are set to $\mu_s = 0.1$ cm^{-1}, $\mu_{a,0} = 0.01$ cm^{-1} for all the tests, and Gaussian noise at 1% level is added to the logarithmic fluctuation to simulate uncertainties in the data and the real behaviour of technical equipment. Three different cases are generated

A) a single circular inclusion is placed at $(x, y) = (0, 2)$ cm with radius .3 cm. The absorption coefficient of the inclusion is set to $\mu_{a,\text{inc}} = 2\mu_{a,0}$

B) two circular inclusions are placed at $(x, y) = (\pm2, 2)$ cm both with radius 0.3 cm. The absorption coefficient of the inclusions is set to $\mu_{a,\text{inc}} = 2\mu_{a,0}$

C) as in case B) but the absorption coefficient of the left inclusion is set to $\mu_{a,\text{inc}} = 5\mu_{a,0}$, while the right inclusion has $\mu_{a,\text{inc}} = 2\mu_{a,0}$.

Tests are performed using the elastic-net regularization method implemented in the `glmnet` software package [16] and alternatively using the Bregman procedure. The results are depicted in Fig. 2, where the two approaches are compared with the state–of–the–art techniques.

As for the elastic-net approach, it allows to correctly locate the inclusion(s) with their absorption coefficient for tests A) and B). In test C) the localization is less precise, but the expected values of the absorption coefficient are well estimated. As for the Bregman procedure, choosing the ridge regression reveals again to be not a very suitable option: although this procedure increases the quality of the reconstruction by the contrast enhancement ([4, 5]), the inclusion is not clearly identified and the recovered halo can be confused with some artefacts. On the other hand, setting the LASSO regularization really benefits from the Bregman implementation: the results for test (A) show that the location of the inclusion is properly recognized and its absorption value μ_a is correctly reconstructed. The reconstruction of test B)
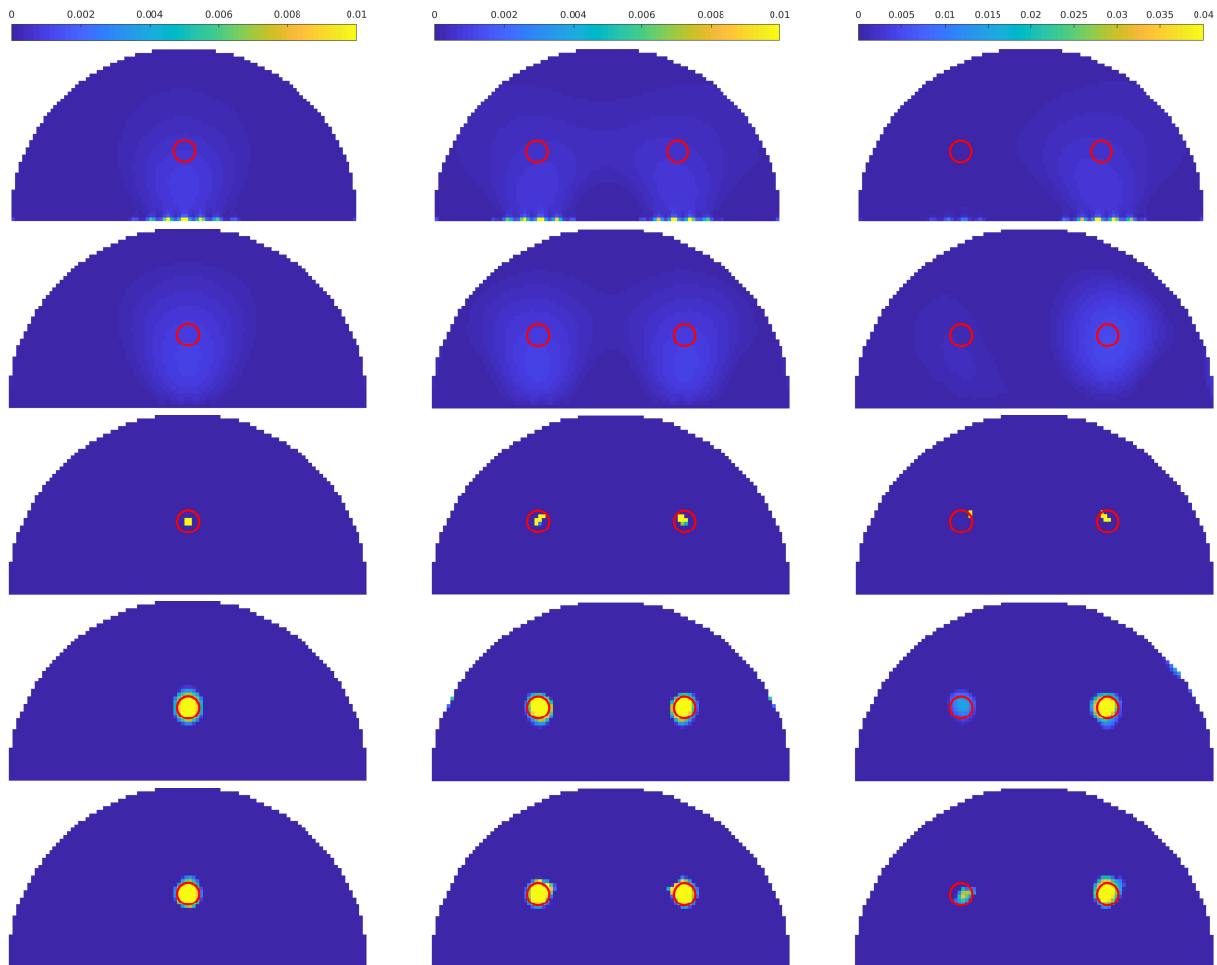
Figure 2. Results for test cases A, B, C in column–wise order. Top to bottom: Tikhonov regularization, Bregman procedure coupled with ℓ_2–norm, pure LASSO approach, Bregman coupled with ℓ_1–norm and eventually elastic–net procedure with $\alpha = 0.5$.

suffers from a small displacement, but the value of μ_a is well recovered. Finally, in test C) both inclusions are not completely placed in the correct location, but the absorption value is close to the correct one. This misplacement is due to the different values of absorption coefficient, hence the regularization parameter promotes the higher one.

5. Conclusions
This work has investigated two original approaches to regularize the inverse problem arising in DOT reconstruction: the elastic–net regularization and the Bregman procedure. The experimental results attest that both techniques actually improve the quality of the estimation of the distribution coefficient, both in localizing the inclusion and in recovering its correct value. Future work will involve the combination of the two procedure and the employment of local and adaptive regularization techniques.

References
[1] Arridge S R 1999 Optical tomography in medical imaging *Inv. Prob* **15** R41–R93
[2] Arridge S R and Schotland J C 2009 Optical tomography: forward and inverse problems *Inv. Prob* **25**, 123010

[3] Banerjee O, El Ghaoui L and d'Aspremont A 2008 Model Selection Through Sparse Maximum Likelihood Estimation for Multivariate Gaussian or Binary Data *J. Mach. Learn. Res.* **9** 485–516

[4] Benfenati A and Ruggiero V 2013 Inexact Bregman iteration with an application to Poisson data reconstruction *Inv. Prob* **29** 065016

[5] Benfenati A and Ruggiero V 2015 Image regularization for Poisson data *J. Phys.: Conf. Series* **657** 012011

[6] Benfenati A and Ruggiero V 2015 Inexact Bregman iteration for deconvolution of superimposed extended and point sources *Comm. Nonlin. Sci. Num. Sim.* **21** 210–24

[7] Benfenati A, La Camera A and Carbillet M 2016 Deconvolution of post-adaptive optics images of faint circumstellar environments by means of the inexact Bregman procedure *A&A* **586** A16

[8] Bonettini S, Benfenati A and Ruggiero, V 2016 Scaling techniques for ε–subgradient methods, *SIAM J. Optim.* **26**, 1741–1772

[9] Benfenati A, Chouzenoux E and Pesquet J–C 2018 A Nonconvex Variational Approach for Robust Graphical LASSO 2018 *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2018)* 3969–73

[10] Boas D A, Brooks D H, Miller E L, DiMarzio C A, Kilmer M and Gaudette R J 2001 Imaging the body with diffuse optical tomography *IEEE Signal Process. Mag.* **18** 57–75

[11] Causin P, Naldi G and Weishaeupl R–M 2019 Elastic Net Regularization in Diffuse Optical Tomography Applications *IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)* to appear

[12] Dehghani H, Pogue B W, Poplack S P and Paulsen K D 2003 Multiwavelength three-dimensional near-infrared tomography of the breast: initial simulation, phantom, and clinical results *Appl Opt* **42** 135–45

[13] De Mol C, Mosci S, Traskine M and Verri A 2009 A regularized method for selecting nested groups of relevant genes from microarray data *J. Comput. Biol.* **16** 677–90

[14] Durduran T, Choe R, Baker W B and Yodh A G 2010 Diffuse optics for tissue monitoring and tomography *Rep. Prog. Phys.* **73** 076701

[15] Friedman J, Hastie T and Tibshirani R 2008 Sparse inverse covariance estimation with the graphical LASSO *Biostatistics* **9** 432–41

[16] J. Friedman J, Hastie T and Tibshirani R 2010 Regularization paths for generalized linear models via coordinate descent *J. Stat. Softw.* **33** 1–22

[17] Gaudette R J, Brooks D H, DiMarzio C A, Kilmer M E, Miller E L, Gaudette T and Boas D A 2000 A comparison study of linear reconstruction techniques for diffuse optical tomographic imaging of absorption coefficient *Phys. Med. Biol.* **45**, 1051–70

[18] Hansen P C 2001 The L-curve and its use in the numerical treatment of inverse problems *Computational Inverse Problems in Electrocardiology* **4** 119–42

[19] Hoshi Y and Yamada Y 2016 Overview of diffuse optical tomography and its clinical applications *J. Biomed. Opt.* **21** 1–11

[20] Ishimaru A 1978 Wave propagation and scattering in random media vol **2** (Academic Press New York)

[21] Chengcai L, Dongdong Y, Shuang Z, Yu A and Yifang H 2015 Reconstruction Method for Optical Tomography Based on the Linearized Bregman Iteration with Sparse Regularization *Comp. Math. Meth. Med.* **2015**

[22] Meinshausen N and Bühlmann P 2006 High-Dimensional Graphs and Variable Selection with the Lasso *Ann. Statist.* **34** 1436–62

[23] Prieto K and Dorn O 2016 Sparsity and level set regularization for diffuse optical tomography using a transport model in 2D *Inv. Prob.* **33** 014001

[24] Pogue B W, Poplack S P, McBride T O , Wells W A, Osterman K S, Osterberg U L and Paulsen K D 2001 Quantitative Hemoglobin Tomography with Diffuse Near-Infrared Spectroscopy: Pilot Results in the Breast *Radiology* **218** 261–66

[25] Sun Z, Wang Y, Jia K and Feng J 2016 Comprehensive study of methods for automatic choice of regularization parameter for diffuse optical tomography *Opt. Eng.* **56** 1–10

[26] Süzen M, Giannoula A and Durduran T 2010 Compressed sensing in diffuse optical tomography *Opt. express* **18** 23676–90

[27] Tibshirani R 1996 Regression shrinkage and selection via the LASSO *J. R. Stat. Soc. B* 267–88

[28] Yalavarthy P K, Pogue B W, Dehghani H, Carpenter C M, Jiang S and Paulsen K D, Structural information within regularization matrices improves near infrared diffuse optical tomography *Opt. Express* **15** 8043–58

[29] Zou H and Hastie T 2005 Regularization and variable selection via the elastic net *J. R. Stat. Soc. B* **67** 301–20

[30] Zhang X, Burger M and Osher S 2011 A Unified Primal-Dual Algorithm Framework Based on Bregman Iteration, *J. Sci. Comput.* **46** 20–46