D04

Proceedings of the 5th International Conference on Inverse Problems in Engineering: Theory and Practice, Cambridge, UK, 11-15th July 2005

DIFFUSION REGULARISATION METHODS OF THE NON-LINEAR INVERSE PROBLEM FOR DIFFUSE OPTICAL TOMOGRAPHY

A. Douiri, M. Schweiger, J. Riley and S. R. Arridge
Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK
e-mail: {a.douiri,m.schweiger,j.riley,s.arridge}@cs.ucl.ac.uk

Abstract - Diffusion Optical Tomography (DOT) is a non-invasive functional imaging modality that aims to image the optical properties of biological organs. The forward problem of the light propagation of DOT can be modelled as a diffusion process and is expressed as a differential diffusion equation with boundary conditions. This model can be considered as the form of a general large scale non-linear inverse problem. In this paper, we formulate a solution to the DOT inverse problem as a minimisation of some functional energy of the solution and the data. One component of this functional is a measure of the discrepancy between the measured data and the data produced by representation of the modelled object. Minimisation of this term alone would force the solution to be consistent with the data and the solver used for this purpose. But in practice data is always accompanied with noise and with some jump discontinuities, these will unavoidably yield unsatisfactory solutions, so some regularisation to restore the solution is required. We add a diffusion regularisation potential with a statistical threshold to the objective functional. This term aims to reduce the noise associated with the data and to preserve the edges in the solution by a diffusion flux analysis of the local structures in the object. To accelerate the iterative solver we introduce a particular method called the Lagged Diffusivity Newton-Krylov method. The whole proposed strategy, which makes use of an a prior diffusion information and an specific algorithm, has been developed and evaluated on simulated data.

1. DIFFUSE OPTICAL TOMOGRAPHY

Diffusion Optical Tomography (DOT) is an emerging medical imaging modality that aims to image the optical properties of biological tissue, particularly the peripheral muscle, breast and the brain [12, 3, 18, 7]. It is non-invasive, portable and can produce images of clinically relevant parameters which cannot be obtained by other diagnostic methods, such as blood volume and oxygenation.

The data acquisition systems for optical tomography consist of a light source as an infrared laser, which is used to illuminate the body surface at different source locations. The light which has propagated through the volume under investigation is measured at multiple detectors locations on the surface. The intensity and the path-length distribution of the detected light provides information about the optical properties of the transilluminated tissue. This boundary data measurement can be used to recover the spatial distribution of internal absorption and scattering coefficients.

The forward problem of the light propagation of DOT can be modelled as a diffusion process [1]. Mathematically, this is expressed by the diffusion equation, which is given in the frequency domain by, $r \in \Omega \subset \mathbb{R}^N (N = 2, 3)$:

$$\left(-\nabla \cdot \kappa \left(r \right) \nabla + \mu_a \left(r \right) + \frac{i\omega}{c} \right) \Phi \left(r, \omega \right) = q \left(r, \omega \right) \tag{1}$$

with Robin-type boundary condition, $m \in \partial\Omega$:

$$\Phi \left(m, \omega \right) + 2\kappa(m) A \frac{\partial \Phi \left(m, \omega \right)}{\partial n} = 0 \tag{2}$$

where ω is the frequency parameter, Φ is the photon density, c is the velocity of light, q is an isotropic source of light in the medium, A is a boundary term which incorporates the refractive index mismatch at the tissue-air boundary, n is the outward normal at $\partial\Omega$, κ and μ_a are the diffusion and absorption coefficients, respectively.

We define here:

$$\kappa = \frac{1}{3 \left(\mu_a + \mu_s \right)} \tag{3}$$

where μ_s is the reduced coefficient.

The measurable quantity is given by the boundary exitance:

$$\Gamma \left(m, \omega \right) = -\kappa(m) \frac{\partial \Phi \left(m, \omega \right)}{\partial n} \tag{4}$$

By using the boundary condition (2), we have:

$$\Gamma = \frac{1}{2A}\Phi \tag{5}$$

The data acquisition provides the exitance Γ and expressed in a complex polar form $|\Gamma|\, e^{i \arg \Gamma}$, where $\arg \Gamma$ is the phase shift and $|\Gamma|$ is the modulation amplitude. The measurement appropriate for the reconstruction is given by the logarithm of the polar form of the exitance.

2. INVERSE PROBLEM AND REGULARISATION

The forward problem can be stated as:
Given a distribution of light sources q on the boundary $\partial\Omega$ of a domain Ω, and a distribution of tissue optical parameters $x = (\mu_a, \mu_s)$ within Ω, find the resulting measurement set y on $\partial\Omega$.
and the inverse problem:
Given a distribution of light sources q on the boundary $\partial\Omega$ and data y on $\partial\Omega$, find $x = (\mu_a, \mu_s)$ in Ω.

Consider S source positions $q_j\,(j = 1, ..., S)$ at the boundary $\partial\Omega$ and M_j measurement positions $m_{i(j)} \in \partial\Omega$, $(i = 1, ..., M_j)$ for the jth source q_j, resulting in a total number of measurements $M_{TOT} = \sum_{j=1}^{S} M_j$. Then, the forward problem can be expressed by the form of a nonlinear operator and is presented by:

$$F(x) = y \tag{6}$$

where the nonlinear forward operator $F : H^1(\Omega) \times H^1(\Omega) \rightarrow H^{-\frac{1}{2}}(\partial\Omega)$ maps a distribution of the object $x = (\mu_a, \mu_s) \in H^1(\Omega) \times H^1(\Omega)$ into a data image $y \in H^{-\frac{1}{2}}(\partial\Omega)$.

The measurement set y is given by:

$$y = \left[\left(\arg \Gamma, \log\left(|\Gamma|\right) \right)_{i(j)} \right]_{\substack{1 \leq i \leq M_j \\ 1 \leq j \leq S}} \tag{7}$$

where the tuple $y_{i(j)} = (\arg \Gamma, |\Gamma|)_{i(j)}$ represents the measurement data collected at the boundary site $m_{i(j)} \in \partial\Omega$ associated to the source distributions q_j.

The inverse problem consists in finding an estimate for the object x which is mapped by the operator F to a perfect data $\bar{y}$. The measured data y approximates the perfect data $\bar{y}$ with:

$$\|y - \bar{y}\| \leq \delta, \; where \; \delta > 0. \tag{8}$$

To solve the inverse problem (6), we had to check some assumptions in the forward operator F. By Hadamard definition [10], the problem (6) is well-posed if F is a homeomorphism between the complete metric spaces $H^1(\Omega) \times H^1(\Omega)$ and $H^{-\frac{1}{2}}(\partial\Omega)$, i.e. the solution of (6) is unique (F is injective), for every $y \in \mathbb{R}^M$ there exists a solution (F is surjective) and the solution of (6) depends continuously on the data y (F is continuously invertible). The problem is ill-posed if it is not well-posed. A mathematically well-posed problem may be ill-posed in practice, the injectivity, subjectivity and stability can not be ensured after the postprocessing stage. The approximation of the forward model, sampling the data, discretisation and other fluctuations may misrepresent the aspect of the problem. The data y may not belong to the range(F) (F is not surjective) or/and more than one vector in the subspace $F^{-1}(\{y\})$ may be solutions of the problem (F is not injective). Hence, it seems natural to replace the data y by its projections onto range(F) to reassure the subjectivity. Moreover, we can select a meaningful solution in the subset $F^{-1}(\{y\})$ by introducing a prior information which involves the choice of some inference principle as smoothness of the solution or other restoration criteria to reduce the subset $F^{-1}(\{y\})$ (injectivity condition) and regularize the inverse problem. Furthermore, the solution may be very sensitive to small perturbations of the data (or measures). A small perturbation δy may give rise to large perturbations $\delta x = F^{-1}(\delta y)$, which may have dramatic consequences on the interpretation of the solution. In this case as well, the constraints need to be relaxed. The combination of these strategies for solving the problem (6) leads to recover a solution that optimize a prior function $\Psi(x)$ over the subspace $F^{-1}(\{\tilde{y}\})$, where $\tilde{y}$ is the projection of y onto the range of F. This can be expressed by:

$$(\mathcal{P}) \begin{cases} Minimize & \Psi(x) \\ subject\ to: & x \in F^{-1}(\{\tilde{y}\}) \end{cases} \tag{9}$$

A standard and equivalent strategy of the problem (9) is to reformulate it to a minimisation system:

$$(\mathcal{P}) \begin{cases} \tilde{x} = \arg\min\limits_{x} \Psi(x) & \text{regularity term} \\ \tilde{x} = \arg\min\limits_{x} D(y, F(x)) & \text{fidelity term} \end{cases} \tag{10}$$

where D is a distance between $F(x)$ and y. The fidelity or fit term ($D(y, F(x))$) can be regarded as a tolerant expression of the inverse problem. We can replace this distance by some other attracting function y to $F(x)$.

A example of a prior function $\Psi(x)$ also called regularity term and a distance D can be chosen as:

- $\Psi(x) = \|x\|_2^2$, minimum norm or $\Psi(x) = \int x \ln x$, maximum entropy.
- $D(F(x), y) = \|y - F(x)\|_p^p$, L^p, $1 \leq p < \infty$ distance.

This strategy leads us to the minimisation problem formulated as follow:

$$(\mathcal{P}) \qquad \textit{Minimize} \quad E(x) := D(y, F(x)) + \tau\Psi(x) \tag{11}$$

where $\tau > 0$ is a regularisation parameter to be prescribed.

Figure 1 summarizes the common strategy discussed above and used for many inverse problem applications.

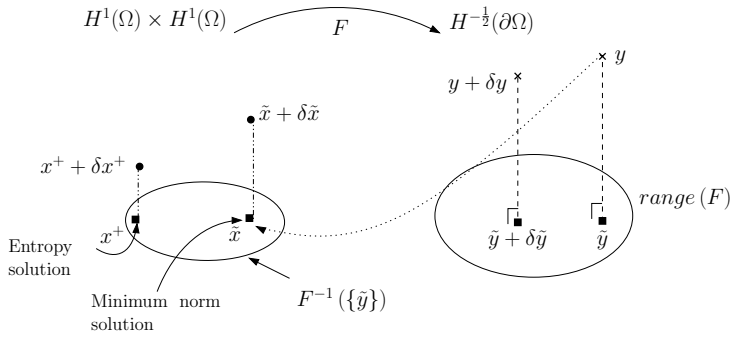


Figure 1: Summary of the strategy adopted for inverse problem.

In this paper, we will study precisely the influence of the regularity term such that the problem has a unique and stable solution, that is physically relevant and computable. We consider that:

$$\Psi(x) = \int \psi(|\nabla x|) \tag{12}$$

where $\psi(|\nabla x|)$ is scalar function whose Gradient $\psi'(|\nabla x|)$ is the flux function and $\frac{\psi'(|\nabla x|)}{|\nabla x|}$ is the diffusion function.

Under the usual assumptions of a system corrupted by multivariate Gaussian additive noise and according to the maximum likelihood principle, a solution of the DOT inverse problem can be given by minimising the following objective function over the domain $H^1(\Omega) \times H^1(\Omega)$ of the object $x = (\mu_a, \mu_s)$:

$$(\mathcal{P}) \qquad \textit{Minimize} \quad E(x) := \frac{1}{2}\|y - F(x)\|_{\mathsf{R}} + \tau \int_\Omega \psi(|\nabla x|) \tag{13}$$

where x is the original object to be recovered, R is the data-space correlation matrix and $\|.\|_{\mathsf{R}}$ is the data-space norm.

The behaviour of the function ψ, influence the local direction of the regularisation. A specific estimation of this function in the minimisation problem is crucial. In this work we will suggest a methodology for choosing an adequate function ψ that could be used in a large scale nonlinear inverse problem, particulary in DOT.

Hypothesis 1. *Assume that ψ verifies the following hypothesis:*

1. *ψ is twice continuously differentiable.*
2. *$\psi(0) = 0$.*
3. *$\psi(s)$ is increasing function for all $s \geq 0$.*

4. *$\psi(s)$ is convex for all $s \geq 0$.*

5. *$\psi(\sqrt{s})$ is concave for all $s \geq 0$.*

We denote by:

$$E(x) = E_{LS}(x) + \tau E_{Reg}(x) \tag{14}$$

$$E_{LS}(x) = \frac{1}{2}\|y - F(x)\|_{\mathsf{R}} \tag{15}$$

$$E_{Reg}(x) = \int_{\Omega} \psi(|\nabla x|) \tag{16}$$

The solution is given by:

$$x_{solution} = \arg\min_{x}\{E(x)\} \tag{17}$$

3. THE LAGGED DIFFUSIVITY NEWTON-KRYLOV SOLVER

Let x be an estimate for the minimizer of E and assume that $E(x)$ is twice Fréchet differentiable at x. Newton's method consists of expanding the functional $E(x)$ in the second order at x, the quadratic approximation to $E(x+h)$ is:

$$Q(h) = E(x) + E'(x) \cdot h + \frac{1}{2}(E''(x) \cdot h) \cdot h \tag{18}$$

$E'(x)$ and $E''(x)$ denote respectively the Gradient and the Hessian of the functional $E(x)$.

Consider that $E''(x)$ is positive definite, then $Q(h)$ has a unique minimizer which satisfies:

$$E'(x) + E''(x) \cdot h = 0 \tag{19}$$

If we replace the estimation x by x_k and $x+h$ by $x_{k+1} := x_k + h_k$ as the new estimate for the minimizer E, where h_k is a solution of the linear system $E'(x_k) + E''(x_k) \cdot h_k = 0$, we obtain the Newton iteration:

$$x_{k+1} = x_k - [E''(x_k)]^{-1} E'(x_k) \tag{20}$$

The are two strategies to improve the convergence of the Newton method: the line search approach called Damped-Newton algorithm and the trust region called Levenberg-Marquardt algorithm approach [16]. In this work, we will use the Levenberg-Marquardt method. Moreover, to avoid the explicit formation of the whole Hessian and the problem of lack of memory due to the large-scale of the DOT, we will use the generalised minimal residual method (GMRES) in Krylov subspace for the linear stage [20, 9], which evaluates and stores sequentially only the result of matrix-vector operation. The algorithm is as follow:

$$\begin{array}{l} \text{1) initialisation} \\ \quad x_0 = \text{initial guess} \\ \text{2) iterations: for } k = 0, 1, 2, \ldots \text{ do} \\ \quad E_k = E(x_k),\ g_k = E'(x_k),\ H_k = E''(x_k),\ Q_k(h) = E_k + g_k \cdot h + \frac{1}{2}(H_k \cdot h) \cdot h \\ \quad s_k = H_k^{-1} g_k \text{ (Krylov/GMRES methods)} \\ \quad x_{k+1} = x_k - s_k \end{array} \tag{21}$$

To calculate the Gradient of the objective function, we simply compute its Fréchet derivative. We have then that the Gradient of $E(x)$ is as follows:

$$\operatorname{grad} E(x) = \mathcal{J}^T(x)\mathsf{R}(F(x) - y) + \tau \mathcal{L}(x)x \tag{22}$$

where:

$$\begin{aligned} \mathcal{J}^T(x)h &= F'(x)^T h \\ \mathcal{L}(x)h &= -\nabla \cdot \left(\frac{\psi'(|\nabla x|)}{|\nabla x|} \nabla x \right) \end{aligned} \tag{23}$$

To deal with the non-differentiability of the absolute value $|\nabla x|$ at t=0, many researchers replace $|\nabla x|$ by other approximation functions such as $|\nabla x| = \sqrt{|\nabla x|^2 + \varepsilon}$, $\varepsilon > 0$. In this work, we will choose

a quadratic approximation of ψ at 0 to avoid this problem. We will explain in detail this approximation in the next paragraph.

For the Hessian, we compute the second Fréchet derivative of the functional $E(x)$. We find that the Hessian of $E(x)$ is given by:

$$\operatorname{Hess} E(x) = \left[\mathcal{J}^T(x)\,\mathsf{R}\mathcal{J}(x) + (F(x)-y)^T\,\mathcal{H}(x)\,\mathsf{R}\right] + \tau\left[\mathcal{L}(x) + \mathcal{L}'(x)\,x\right]$$

where:

$$\begin{aligned}\mathcal{H}(x)\,H &= F''(x)\,H\\ \mathcal{L}'(x)\,x\cdot h &= -\nabla\cdot\left(\left(\psi''(|\nabla x|) - \frac{\psi'(|\nabla x|)}{|\nabla x|}\right)\frac{\nabla x\otimes\nabla x}{|\nabla x|^2}\nabla h\right)\end{aligned}$$

where $\otimes$ denotes the Kronecker product.

Given the size of the data, an explicit discretisation of the whole Hessian may exceed the allocated memory. One possible solution is to ignore the useless terms and reformulate this huge Hessian by a smallest and nearest approximation.

On the one hand $\mathcal{H}(x)\mathsf{R}$ is bounded:

$$(F(x)-y)^T\,\mathcal{H}(x)\,\mathsf{R} \le const\,\|y-F(x)\| \tag{24}$$

That means when $\|y-F(x)\| \mapsto 0$ implies that $(y-F(x))^T\,\mathcal{H}(x)\,\mathsf{R}$ is very small. Then, the terms of with $\mathcal{H}(x)\mathsf{R}$ can be ignored in line search or trust region procedure of Newton method. Hence, the term $(F(x)-y)^T\mathcal{H}(x)\mathsf{R}$ can be omitted in the Hessian term. Note that for Levenberg-Marquardt algorithm we add a relaxation term γI, where γ is a positive parameter that will be adjusted at each iteration in respect of the algorithm (21) [6, 15]. The approximated Hessian of the least squares functional becomes:

$$\operatorname{Hess} E_{LS}(x) = \mathcal{J}^T(x)\,\mathsf{R}\mathcal{J}(x) \tag{25}$$

On the other hand, we have :

$$E''_{Reg}(x)(h,h) = \frac{\psi'(|\nabla x|)}{|\nabla x|}\underbrace{\left(|\nabla h|^2 - \frac{(\nabla x\cdot\nabla h)^2}{|\nabla x|^2}\right)}_{\geq 0\ By\ Cauchy\text{-}Schwartz} + \psi''(|\nabla x|)\,\frac{(\nabla x\cdot\nabla h)^2}{|\nabla x|^2} \geq 0 \tag{26}$$

Under the Hypothesis 1 on ψ and if $\mathcal{J}^T\mathsf{R}\mathcal{J}$ is positive definite, it is quite easy to prove the global convergence of $E(x)$, so no trust region strategy is necessary.

However, we have:

$$\mathcal{L}'(x)(h,h) = \underbrace{\left(\psi''(|\nabla x|) - \frac{\psi'(|\nabla x|)}{|\nabla x|}\right)}_{\leq 0\ By\ Concavity\ of\ \psi(\sqrt{s})\ on\ s}\ \frac{(\nabla x\cdot\nabla h)^2}{|\nabla x|^2} \tag{27}$$

So, by dropping the negative semidefinite operator $\mathcal{L}'(x)$ from the Hessian, the Levenberg-Marquardt method still converges to a unique solution. This approximation can be inspired from Gauss-Newton approximation of the Hessian neglecting the term $\mathcal{L}'(x)$ [24].

The actual approximated Hessian becomes:

$$\operatorname{Hess} E_{Reg}(x) = \nabla\cdot\left(\frac{\psi'(|\nabla x|)}{|\nabla x|}\nabla\right) = \mathcal{L}(x) \tag{28}$$

This method is often called the Lagged Diffusivity approach, as the Gradient and the Hessian in Newton algorithm are calculated from the diffusion operator $\mathcal{L}(x)$.

We conclude that the approximated Hessian becomes:

$$\operatorname{Hess} E(x) = \mathcal{J}^T(x)\,\mathsf{R}\mathcal{J}(x) + \tau\mathcal{L}(x) \tag{29}$$

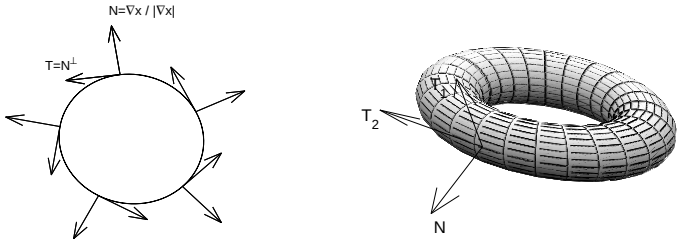


Figure 2: Geometrical interpretation: Left graph is the 2D case and the right graph is the 3D case.

4. ISOTROPIC AND ANISOTROPIC REGULARISATION

Weickert [25] makes an interesting analogy between the physical background of diffusion and its application in image processing. Indeed this inference allows a good understanding of the diffusion process. In physics, the diffusion is a process of equilibration of concentration differences with mass conservation. The isotropic diffusion is a property that preserves its value when measured in different directions. The anisotropic diffusion is a property that has a different value when measured in different directions.

In image processing the concentration could be described by the grey value or the colour index of the image. The diffusion process can be characterized by the behaviour of ψ with respect to the local image structure x. The diffusion function can be adapted to the local image structure. If the diffusion function is constant over the whole image domain, the process is called isotropic or homogeneous diffusion, and if the diffusion function is adapted to the local image structure, we have an anisotropic or nonhomogeneous diffusion process. The anisotropic diffusion model in image processing was introduced by Perona and Malik [17] by using a diffusion function dependent upon the norm of the gradient of the image. In the same context, Cottet and Germain [5] propose another alternative by choosing the diffusion function as a tensor-matrix. A drawback of the structure of the anisotropic diffusion matrix is the increasing of calculation time which discourage its utilisation for the large-scale inverse problems.

To investigate the behaviour of diffusivity, we decompose the Gradient of the regularity term $\nabla.\left(\frac{\psi'(|\nabla x|)}{|\nabla x|}\nabla x\right)$ by using the local object structures, that is, the tangent and the normal directions to the isosurface (isophote lines in 2D). We denote by $N = \frac{\nabla x}{|\nabla x|}$ where $|\nabla x| \neq 0$ is the unit vector in the direction of the Gradient of the object and $T = N^{\perp}$ the hyperplane tangent to the local isosurface. This decomposition can be expressed on the basis of eigenvectors as follows [22]:

Proposition 1.

$$\begin{aligned}\nabla.\left(\frac{\psi'(|\nabla x|)}{|\nabla x|}\nabla x\right) &\overset{3D}{=} \frac{\psi'(|\nabla x|)}{|\nabla x|}\left(x_{T_2T_2} + x_{T_1T_1}\right) + \psi''(|\nabla x|)\, x_{NN} \\ \nabla.\left(\frac{\psi'(|\nabla x|)}{|\nabla x|}\nabla x\right) &\overset{2D}{=} \frac{\psi'(|\nabla x|)}{|\nabla x|} x_{TT} + \psi''(|\nabla x|)\, x_{NN}\end{aligned} \tag{30}$$

where (N, T_1, T_2) (resp (N, T)) is an orthogonal basis in $\mathbb{R}^3$ (resp $\mathbb{R}^2$) (see Figure 2).

(N, T_1, T_2) are the eigenvectors of PHP with $P = I - NN^T$ and $H = \text{Hess}(x)$. For the 2D case, x_{NN} and x_{TT} are the second derivatives of x in the T-direction and N-direction, respectively. And $\left(\sum_{i=1}^{n} -|\nabla x|\,\kappa_i\right) = \Delta x - x_{NN}$. This proposition allows us to understand the action of the regularisation divergence operator in the directions N and $T = N^{\perp}$. That allow us to interpret the diffusion process along the flow line of the associated eigenvector fields.

In the case of isotropic diffusion regularisation, that is Tikhonov's regularisation [23], we have:

$$\Psi(|\nabla x|) = \int \frac{1}{2}|\nabla x|^2 \tag{31}$$

The Gradient of this regularity term is:

$$\nabla \cdot \left(\frac{\psi'(|\nabla x|)}{|\nabla x|}\nabla x\right) = x_{NN} + x_{T_1T_1} + x_{T_2T_2} \tag{32}$$

The isotropic diffusion (Tikhonov) regularisation acts in all directions. Consequently, the smoothing process does not only reduce the noise, but also blurs important structure features such as the edges and contours, thus making them hard to identify. Furthermore, isotropic diffusion regularisation could dislocate structure information when mapping from finer to coarser grids, that is the case in inverse problem solvers which use multigrid methods or mapping from finer mesh to a regular coarse grid.

While the isotropic diffusion (Tikhonov) regularisation does not preserve structures, we need to find a new function ψ that verifies the convergence hypothesis (Hypothesis 1) and allows an efficient noise reduction, taking fully into account the restoration and enhancement of the local structures information present in the object. A good diffusion regularisation consists in the following conditions:

1. The Hypothesis 1 is fulfilled.
2. The regularisation must be isotropic in homogeneous regions, that is when $|\nabla x|$ is small (low gradients).
3. In the vicinity of an edge, the object presents a strong gradient. In order to preserve this edge the diffusion must act in the direction of $N^{\perp}$ and not across it.

The choice of a function satisfying these conditions is based on the directional interpretation, see eqn. (30). These conditions are expressed by the following relations:

1. $\psi \in \mathcal{C}^2(\mathbb{R})$, $\psi(0) = 0$, $\psi'(s) \geq 0$, $\psi''(s) \geq 0$, $(\psi(\sqrt{s}))'' \leq 0$ for all $s \geq 0$.
2. $\psi'(0) = 0$, $\lim\limits_{s \to 0^+} \frac{\psi'(s)}{s} = \lim\limits_{s \to 0^+} \psi''(s) = \psi''(0) > 0$
3. $\lim\limits_{s \to +\infty} \psi''(s) = 0$, $0 < \lim\limits_{s \to +\infty} \frac{\psi'(s)}{s} = \beta < \infty$

Unfortunately, the conditions 2 and 3 are incompatible. One possible solution is to use a composite of two functions, the first one isotropic such as Tikhonov functional and the second one an anisotropic functional such as the total variation functional [19]. An alternative choice is also possible and leads to almost the same result, if we choose Perona-Malik functional [17], Green functional [8], or Tukey's biweight functional [4].
Given that must we decide between flat or edges regions, we can use a statistical threshold in the gradient of the object to make the right decision.

In this work, we will suggest and use this function:

$$\psi_\sigma(|\nabla x|) = \begin{cases} \frac{|\nabla x|^2}{2} & if\, |\nabla x| \leq \sigma \\ \frac{\sigma}{\psi'(\sigma)} \left(\psi(|\nabla x|) - \psi(\sigma)\right) + \frac{\sigma^2}{2} & otherwise \end{cases} \tag{33}$$

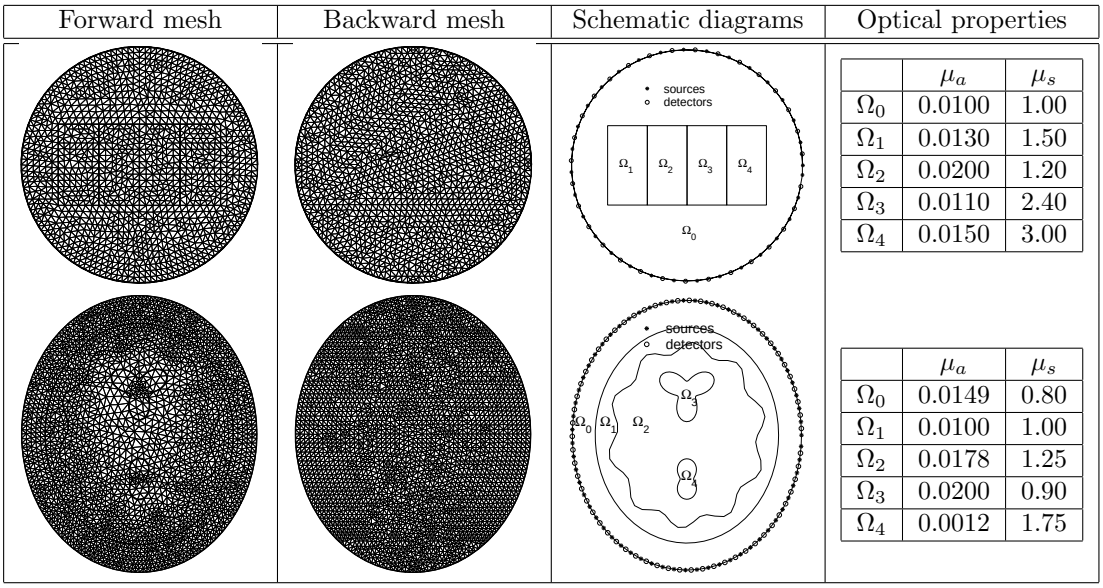
σ is a scale that we adjust in each iteration of the Lagged Diffusivity Newton-Krylov solver. Note here that the choice of σ depends on our hypothesis that we have assumed on ψ as well. An automatic choice of this parameter can be based on the cumulated histogram [17]or on the local geometry [14] of the object at each iteration.
The Gradient of this regularity term is expressed by:

$$\nabla \cdot \left(\frac{\psi'(|\nabla x|)}{|\nabla x|} \nabla x \right) = \begin{cases} x_{NN} + x_{T_1T_1} + x_{T_2T_2} & if\, |\nabla x| \leq \sigma \\ \frac{\sigma}{|\nabla x|} \left(x_{T_1T_1} + x_{T_2T_2}\right) & otherwise \end{cases} \tag{34}$$

For the choice of σ we will use the robust statistic estimator [13], the diffusion can be seen as a robust estimation procedure that estimates a piecewise smooth object from noisy input data. The diffusion function that ensures the isotropy or anisotropic criterion is closely related to the error norm of ψ and the influence function is proportional to the derivative ψ' in the robust statistic analysis.

5. EXPERIMENTS AND RESULTS

We present in this section some results on the comparison of the isotropic diffusion regularisation. The isotropic (Tikhonov) regularisation method and the proposed (adapted diffusion) regularisation method, are tested and compared in simulated 2D and 3D objects. In order to show the advantage of the proposed method compared to the conventional Tikhonov one,particularly, the edges and structures information preserving and enhancement, we consider two 2D phantom objects: a straight edges phantom and a curves phantom with diameter of $70mm$. Each phantom have a scattering and absorption coefficients



	μ_a	μ_s
Ω_0	0.0100	1.00
Ω_1	0.0130	1.50
Ω_2	0.0200	1.20
Ω_3	0.0110	2.40
Ω_4	0.0150	3.00

	μ_a	μ_s
Ω_0	0.0149	0.80
Ω_1	0.0100	1.00
Ω_2	0.0178	1.25
Ω_3	0.0200	0.90
Ω_4	0.0012	1.75

Figure 3: The 2D phantoms: The top row is the straight edges phantom and the bottom row is the curves phantom.

distributions μ_a and μ_s varying relative to each region. 32 sources are placed at equidistance spacing along the surface, and 32 detectors is located between two source sites.The measurements in detectors are acquired from each source except the two closest ones, leading to a total of 960 measurements. We assume that sources to be amplitude-modulated with a frequency of $50MHz$, and each measurement consists of the log of the modulation amplitude and the phase shift [2, 21]. The data y are computed with the finite element method (FEM) of the diffusion forward model using an adapted triangle mesh consisting of 1339 nodes for the straight edges phantom and 2572 for the curves phantom. For the reconstruction we assume no a priori knowledge about the true coefficient distributions and we use homogeneous (constant) values for the initial guess. Starting from this estimate, we perform the reconstruction in a regular grid of 70×70 pixel. The FEM backward basis consist of a regular triangle mesh of 1485 nodes for edges phantom and 3065 nodes for the curves phantom, using linear shape function (P1 approximation). All the data for this 2D simulations was blurred with 1% Gaussian noise and we use the regularisation parameter $\tau = 10^{-5}$ according to the L-curve method [11]. Figure 3 illustrates the two phantom and their optical properties. The reconstructions was performed in respect of our proposed methodology detailed previously in this paper.

Figure 4 shows the reconstructions of the two phantoms with Tikhonov regularisation and the proposed regularisation method. To show the robustness of the proposed method for preserving the structures information we show up the profile lines at different locations. These profiles clearly show the advantage of the proposed regularisation method and prove the inability of the conventional isotropic (Tikhonov) diffusion method. Moreover, the Tikhonov method can change the shape of objects as showed in Figure 4 for the reconstruction of the s of the curved phantoms. To appraise the robustness of the reconstructions of the proposed method reconstruction we compared their H^1-error ($\|x - x_r ec\|_H^1$) with the Tikhonov one. The Table 1 shows the improvement of the new proposed method compared to the Tikhonov method.

	Tikhonov method	Proposed method
Edges phantom	0.27	0.17
Curved phantom	0.96	0.84

Table 1: H^1-error of Tikhonov method and the proposed method.

6. CONCLUSIONS

Visual inspection confirms the hight quality of the anisotropic diffusion regularisation of the inverse problem for DOT. Moreover, anisotropic diffusion regularisation reduces the noise and recovers the edges and structures information present in the object. This improvement in the reconstruction of the optical

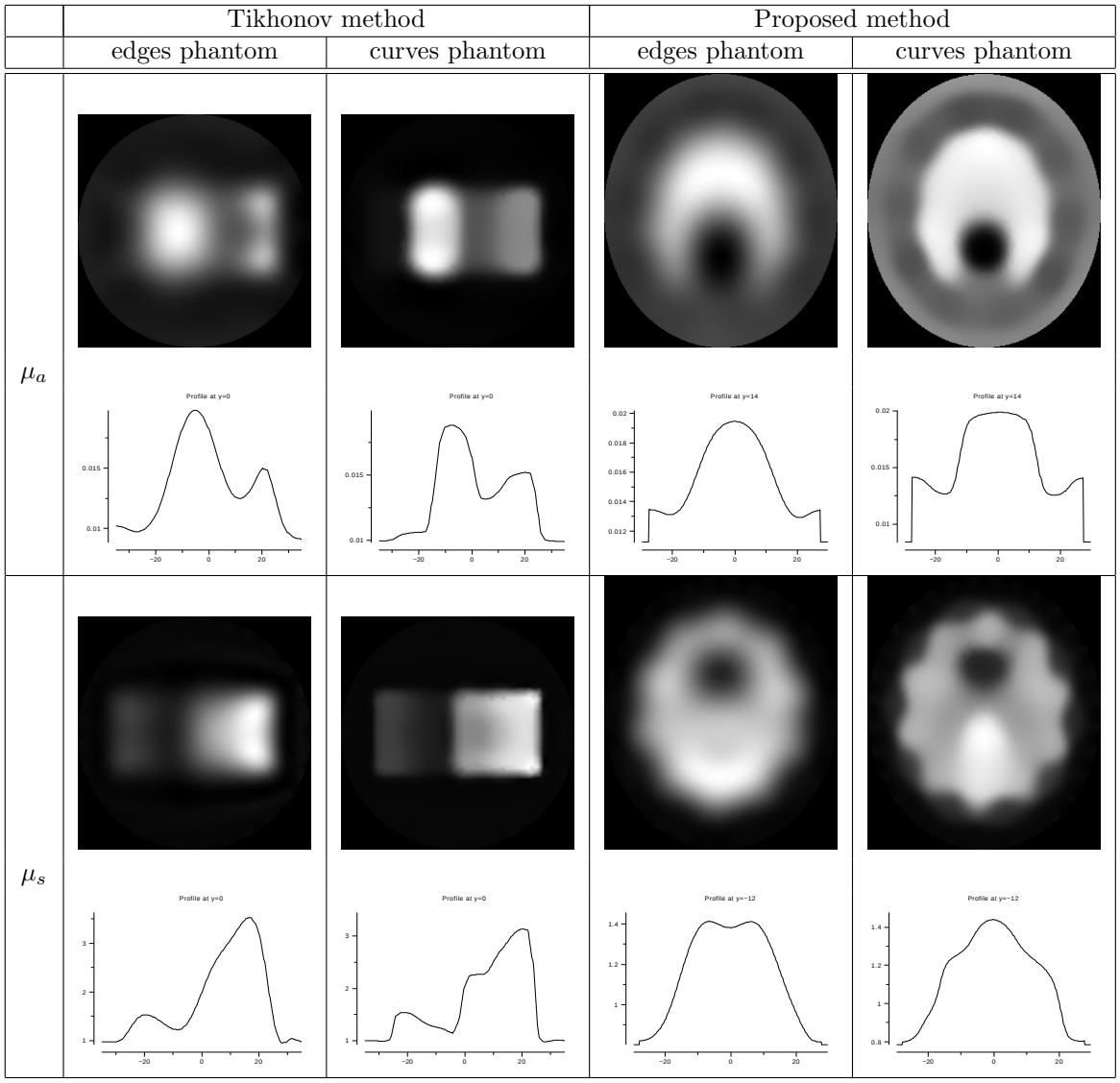


Figure 4: The 2D reconstructions and their profiles. For the edges phantom we have the profiles at y = 0 for a and s and for the curved phantom we have the profile at line y = 14 for μ_a and at line y = -12 for μ_s.

properties with DOT could open a new research development for qualitative and quantitative studies and could gives accurate reconstructed absorption and scattering coefficients that contribute to a better diagnosis in medical optical tomography.

Acknowledgement
The authors would like to acknowledge the financial support received from the EPSRC, GR/R86201/01.

REFERENCES

1. S. R. Arridge, Optical tomography in medical imaging. *Inverse Problems* (1999) **15**(2), R41–R93.

2. S. R. Arridge, M. Schweiger, M. Hiraoka and D. T. Delpy, A finite element approach for modeling photon transport in tissue. *Med. Phys.* (1993) **20**(2), 299–309.

3. S. R. Arridge, P. van der Zee, D. T. Delpy and M. Cope, Reconstruction methods for infrared absorption imaging, *Proc. SPIE Time-Resolved Spectroscopy and Imaging of Tissues*, (eds. B. Chance and A. Katzir) 1991 **1431**, pp. 204–215.

4. M. J. Black, G. Sapiro, D. H. Marimont and D. Heeger, Robust anisotropic diffusion. *IEEE Trans. Image Processing* (1998) **7**(3), 421–432.

5. G. H. Cottet and L. Germain, Image-processing through reaction combined with nonlinear diffusion. *Math. Comput.* (1993) **61**(204), 659–673.

6. R. S. Dembo and T. Steihaug, Truncated-Newton algorithms for large-scale unconstrained optimization. *Mathematical Programming* (1983) **26**, 190–212.

7. S. Fantini, S. A. Walker, M. A. Franceschini, M. Kaschke, P. M. Schlag and K. T. Moesta. Assessment of the size, position, and optical properties of breast tumors in vivo by noninvasive optical methods. *Appl. Opt.* (1998) **37**(10), 1982–1989.

8. P. J. Green, Bayesian reconstructions from emission tomography data using a modified em algorithm. *IEEE Trans. Med. Imaging* (1990) **9**(1), 84–93.

9. A. Greenbaum and L. N. Trefethen, GMRES/CR and Arnoldi/Lanczos as matrix approximation problems. *SIAM J. Matrix Anal. Appl.* (1992) **15**, 778–795.

10. J. Hadamard, Sur les problemes aux derivees partielles et leur signification physique. *Bullentin Princeton University* (1902) **13**, 49–52.

11. P. C. Hansen and D. P. O'Leary, The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM J. Sci. Comput.* (1993) **14**, 1487–1503.

12. J. C. Hebden, F. E. W. Schmidt, M. E. Fry, M. Schweiger, E. M. C. Hillman, D. T. Delpy and S. R. Arridge, Simultaneous reconstruction of absorption and scattering images by multichannel measurement of purely temporal data. *Opt. Lett.* (1999) **24**, 534–536.

13. P. J. Hubert, *Robust Statistics*, Wiley, New York, 1981.

14. D. S. Luo, M. A. King and S. Glick, Local geometry variable conductance diffusion for postreconstruction filtering. *IEEE Trans. Nucl. Sci.* (1994) **41**(6), 2800–2806.

15. E. Mizutani and J. Demmel, On iterative krylov-dogleg trust-region steps for solving neural networks nonlinear least squares problems. *NIPS*, (eds. T. K. Leen, T. G. Dietterich, and V. Tresp), MIT Press, 2000, pp. 605–611.

16. J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer Verlag, New York, 1999.

17. P. Perona and J. Malik, Scale-space and edge-detection using anisotropic diffusion. *IEEE Trans. Patt. Anal. Mach. Int.* (1990) **12**(7), 629–639.

18. B. W. Pogue, K. D. Paulsen, C. Abele and H. Kaufman, Calibration of near-infrared frequency-domain tissue spectroscopy for absolute absorption coefficient quantitation in neonatal head-simulating phantoms. *J. Biomed. Opt.* (2000) **5**(2), 185–193.

19. L. I. Rudin, S. Osher and E. Fatemi, Nonlinear total variation based noise removal algorithm. *Physica D* (1992) **60**, 259–268.

20. Y. Saad and M. H. Schultz, GMRES: A generalized minimal residual algorithm for solving non-symmetric linear systems. *SIAM J. Sci. Statist. Comput.* (1986) **7**(3), 856–869.

21. M. Schweiger, S. R. Arridge, M. Hiraoka and D. T. Delpy, The finite element model for the propagation of light in scattering media: Boundary and source conditions. *Med. Phys.* (1995) **22**(11), 1779–1792.

22. S. Teboul, L. Blanc-Feraud, G. Aubert and M. Barlaud, Variational approach for edge-preserving regularization using coupled pde's. *IEEE Trans. Image Processing* (1998) **7**(3), 387–397.

23. A. Tikhonov, Regularization of incorrectly posed problems. *Soviet Math. Doklady* (1963) **4**, 1624–1627.

24. C. R. Vogel, *Computational Methods for Inverse Problems*, SIAM, Philadelphia, 2002.

25. J. Weickert, A review of nonlinear diffusion filtring, In *Scale-Space Theory in Computer Vision*, Lecture Notes in Comp. Science, **687**, pp. 3–28, Springer-Verlag, Berlin, 1997.