Fluorescence diffuse optical tomography using the split Bregman method

J. F. P.-J. Abascal,[a)] J. Chamorro-Servent, and J Aguirre
Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, 28911 Madrid, Spain

S. Arridge and T. Correia
Department of Computer Science, University College London, Gower Street, London WC1E 6BT, United Kingdom

J. Ripoll
Institute of Electronic Structure and Laser Foundation for Research and Technology Hellas, P.O. Box 1527, 71110 Heraklio, Greece and Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, 28911 Madrid, Spain

J. J. Vaquero
Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, 28911 Madrid, Spain

M. Desco
Departamento de Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, 28911 Madrid, Spain and Unidad de Medicina y Cirugía Experimental, Hospital General Universitario Gregorio Marañón, CIBER de Salud Mental (CIBERSAM), 28007 Madrid, Spain

Purpose: Standard image reconstruction methods for fluorescence Diffuse Optical Tomography (fDOT) generally make use of L2-regularization. A better choice is to replace the L2 by a total variation functional that effectively removes noise while preserving edges. Among the wide range of approaches available, the recently appeared Split Bregman method has been shown to be optimal and efficient. Furthermore, additional constraints can be easily included. We propose the use of the Split Bregman method to solve the image reconstruction problem for fDOT with a nonnegativity constraint that imposes the reconstructed concentration of fluorophore to be positive.
Methods: The proposed method is tested with simulated and experimental data, and results are compared with those yielded by an equivalent unconstrained optimization approach based on Gauss Newton (GN) method, in which the negative part of the solution is projected to zero after each iteration. In addition, the method dependence on the parameters that weigh data fidelity and nonnegativity constraints is analyzed.
Results: Split Bregman yielded a reduction of the solution error norm and a better full width at tenth maximum for simulated data, and higher signal-to-noise ratio for experimental data. It is also shown that it led to an optimum solution independently of the data fidelity parameter, as long as the number of iterations is properly selected, and that there is a linear relation between the number of iterations and the inverse of the data fidelity parameter.
Conclusions: Split Bregman allows the addition of a nonnegativity constraint leading to improve image quality.

Key words: diffuse optical tomography, L1-regularization, total variation, constrained optimization, split Bregman

I. INTRODUCTION

Diffuse optical tomography in fluorescent mode (fDOT) is becoming an important preclinical noninvasive technique for small animal imaging.[1–5] Tomographic algorithms are based on modeling the propagation of light through biological tissue which can be described by the diffusion approximation since most biological tissues are highly-scattering media in the near-infrared range.[6] Solution to tomographic image reconstruction can be tackled as a linear inverse problem of the recovery of the concentration of fluorophore from fluorescent boundary data and is usually addressed by means of implicit regularization, using iterative methods like the algebraic reconstruction technique,[3,7,8] or by unconstrained optimization methods based on L2-regularization.[9,10] While these methods constitute the common choice for practical applications, as they are easy to implement, it is well known that they lead to over-regularized solutions, smoothing out edges in the image. In contrast, the L1-norm of the gradient of the image, the total variation (TV) functional, preserves edges and has been termed as the proper norm for images.[11,12] However, since the TV functional is nonlinear and nondifferentiable, it requires stable and efficient iterative algorithms.

Constrained optimization using conventional algorithms[12,13] has been shown to lead to optimal image reconstruction. Nevertheless, due to its high computational cost and complicated implementation, more simplified approaches are generally adopted. A straightforward approach is to formulate an equivalent unconstrained optimization problem.[13] An alternative approach is to convert the constrained problem into an equivalent unconstrained optimization method and to impose the constraint using an iterative procedure. For example, in the so-called continuation methods the regularization parameter is iteratively increased to enforce the constraint,[13] yet they can be inefficient and unstable. Since novel methods based on the Bregman iteration enforce the constraints in an optimum way, adding the Bregman iteration to the unconstrained problem enforces the solution to converge to the solution of an equivalent constrained problem.[14] Exploiting this idea, the Split Bregman (SB) method solves constrained optimization problems with convex nondifferentiable functionals on an efficient and simple manner.[15,16] For instance, for L1-constraints such as the TV, the SB method decouples L1- and L2-functionals, and minimizes them separately, the L2-part by using conventional methods, and the L1-part by straightforward shrinkage formulas. SB has also a close connection to alternating splitting methods.[17]

TV, the Bregman iteration, and splitting methods have been widely applied to image denoising and magnetic resonance imaging.[11,12,14,16,18] However, very few studies have addressed their use for DOT and fDOT. Earlier implementations of TV for DOT were based on unconstrained methods, such as a regularized least squares approach[19] or a Gauss - Newton (GN) method with anatomical prior.[20] L1 and TV were also applied to bioluminescence tomography using the interior point method.[21,22] Recent studies have applied an iterated shrinkage method[23] and an augmented splitting Lagrangian approach[24] to fDOT simulated data, and a splitting method based on anisotropic diffusion regularization with anatomical prior to fDOT phantom and *ex-vivo* data.[25] Furthermore, the Split Bregman formulation has been employed to enforce a nonnegativity constraint for image denoising with Poissonian statistics.[26]

The aim of this work is to validate the Split Bregman method to minimize the TV of the image with a nonnegativity constraint for fDOT, and to test its performance using simulated and experimental phantom data. Results are compared with an equivalent unconstrained optimization approach solved with a GN method, in terms of the solution error norm and image quality. In addition, the ability to enforce the nonnegativity constraint is analyzed and compared with GN method projecting to zero the negative part of the solution after each iteration.

II. METHODS

II.A. Forward problem

The propagation of light through tissue can be approximated by the diffusion equation, valid for weakly anisotropic and highly-scattering media, for which the reduced scattering coefficient μ_s' is much larger than the absorption coefficient μ_a.[6] Hence, in fluorescence diffuse optical tomography, the excitation photon density $\Phi_{ex}(\mathbf{r})$ and emission photon density $\Phi_{em}(\mathbf{r})$ are given by the solution of a couple of diffusion equations.[9,27] The excitation photon density $\Phi_{ex}(\mathbf{r})$ is emitted by an external source $q_0(\mathbf{r}_s)$ and the emission photon density $\Phi_{em}(\mathbf{r})$ comes from a fluorophore with fluorescent yield $u(\mathbf{r}_{fl})$, which accounts for its quantum efficiency, absorption parameter and concentration of fluorophore. In constant-wave mode (zero frequency), excitation and emission photon densities are given by

$$-\nabla\cdot\kappa(\mathbf{r})\nabla\Phi_{ex}(\mathbf{r})+\mu_a(\mathbf{r})\Phi_{ex}(\mathbf{r})=q_0(\mathbf{r}_s) \quad (1)$$

$$-\nabla\cdot\kappa(\mathbf{r})\nabla\Phi_{em}(\mathbf{r})+\mu_a(\mathbf{r})\Phi_{em}(\mathbf{r})=u(\mathbf{r}_{fl})\Phi_{ex}(\mathbf{r}). \quad (2)$$

where $\kappa(\mathbf{r})=[3(\mu_a(\mathbf{r})+\mu'_s(\mathbf{r}))]^{-1}$ is the diffusion coefficient, and for simplicity it is assumed that optical parameters are equal at excitation and emission (they have close wavelengths) and that the absorption parameter is not influenced by the presence of the fluorophore.

Let's define a Green's function $\mathbf{G}(\mathbf{r},\mathbf{r}')$, for a point source located at $\mathbf{r}'$, that solves

$$[-\nabla\cdot\kappa(\mathbf{r})\nabla+\mu_a(\mathbf{r})]\mathbf{G}(\mathbf{r},\mathbf{r}')=\delta(\mathbf{r}-\mathbf{r}'), \quad (3)$$

then, photon densities that solve Eqs. (1) and (2) are given by

$$\Phi_{ex}(\mathbf{r})=\int d\mathbf{r}'\mathbf{G}(\mathbf{r},\mathbf{r}')q_0(\mathbf{r}') \quad (4)$$

$$\Phi_{em}(\mathbf{r})=\int d\mathbf{r}'\mathbf{G}(\mathbf{r},\mathbf{r}')u(\mathbf{r}')\Phi_{ex}(\mathbf{r}'). \quad (5)$$

The fluorescent data are usually normalized to the excitation data to reduce the dependence on the geometry and optical coefficients.[8,28,29] Hence, the fDOT data measured at the detector position $\mathbf{r}_d$ is defined as

$$g(\mathbf{r}_d)=\frac{\Phi_{em}^{meas}(\mathbf{r}_d)}{\Phi_{ex}^{meas}(\mathbf{r}_d)}. \quad (6)$$

It can be shown that the Jacobian matrix J_{ij} that relates each measurement g_i (the subindex i accounts for each source-detector pair) to the concentration of fluorophore u_j at the element Ω_j of the discretized domain Ω is

$$J_{ij}=\frac{\partial g_i}{\partial u_j}=\frac{1}{\Phi_{ex}^{meas}}\int_{\Omega_j}d\mathbf{r}_j\tilde{\Phi}(\mathbf{r}_j,\mathbf{r}_d)\Phi_{ex}(\mathbf{r}_j,\mathbf{r}_s), \quad (7)$$

where $\tilde{\Phi}(\mathbf{r}_j,\mathbf{r}_d)$ is the adjoint field solved by considering a source at the detector position in Eq. (4). Thus, the fDOT problem can be expressed as a linear system[9,27,30]

$$\mathbf{g}=\mathbf{J}\mathbf{u}, \quad (8)$$

where $\mathbf{J}$ is a $M\times N$-matrix, $\mathbf{u}$ is a $N\times 1$-vector of fluorophore coefficients, and $\mathbf{g}$ is a $M\times 1$-vector, with N the number of voxels and M the number of measurements.

Light propagation was modeled using the TOAST finite element package for diffuse optical tomography[31,32] for both data simulation and building the Jacobian.

II.B. Image reconstruction problem

Let $\Omega\subset\mathbb{R}^3$ be the domain, $\partial\Omega$ the boundary of the domain, $u\in\Omega$ the reconstructed image of concentration of

fluorophore, $g \in \partial\Omega$ the boundary data, and $J : u \mapsto g$ [Eq. (7)], the linear operator that maps image into data; one aims at the recovery of the image u from boundary data g such that the image has a low total variation.

II.B.1. Total variation

Let u be a smooth function defined in a domain Ω, its total variation $R(u)$ is given by

$$R(u) = \|\nabla u\|_1 = \int_\Omega d\mathbf{r}|\nabla u|, \tag{9}$$

where $\|\nabla u\|_1$ is the 1-norm of the gradient of u, which can be understood as the sum over the domain of the absolute variations of the function u.[12] The nondifferentiability of the absolute functional at the origin can be avoided by using an approximation

$$R(u) = \int_\Omega d\mathbf{r}\Psi(|\nabla u|), \tag{10}$$

where the functional $\Psi(|\nabla u|)$ deals with the nondifferentiability; in this work, we have used

$$\Psi(|\nabla u|) = \sqrt{(\nabla_x u)^2 + (\nabla_y u)^2 + (\nabla_z u)^2 + \beta^2} \tag{11}$$

where $\nabla_i = \partial_i$ and β is a small parameter.

Minimization of TV leads to an image with few oscillations while allowing for sharp discontinuities and thus preserving edges. This can be understood from the gradient of $R(u)$, which being proportional to $\nabla \cdot \kappa \nabla u$, where κ is the diffusion function

$$\kappa = \frac{\Psi'(|\nabla u|)}{|\nabla u|}, \tag{12}$$

preserves edges by decaying to zero as the absolute gradient increases. Other approximations to the diffusion function lead to different edge preserving functions, like Huber and Perona-Malik.[20] More details can be found in Refs. 11, 12, 20, and 25.

II.B.2. Unconstrained optimization

It is common to tackle the reconstruction problem using an unconstrained optimization approach

$$\min_{\mathbf{u}} \Theta(u) = \min_{\mathbf{u}} R(\mathbf{u}) + \frac{\lambda}{2}\|\mathbf{J}\mathbf{u} - \mathbf{g}\|_2^2, \tag{13}$$

where λ is a parameter that weights the data constraint (this would be the reciprocal of the commonly named regularization parameter). Using a gradient-based approach, the solution at step $k+1$ is updated as

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \mathbf{p}^k. \tag{14}$$

An efficient minimization direction $\mathbf{p}$ is given by the GN method[33]

$$\mathbf{p} = -\frac{\Theta'}{\Theta''}, \tag{15}$$

where Θ' stands for the gradient and Θ'' for the Hessian.

II.B.3. Constrained optimization

The constrained inverse problem consists in the minimization of the total variation of the image $R(\mathbf{u})$ subject to a data constraint

$$\min_{\mathbf{u}} R(\mathbf{u}) \quad \text{such that} \quad \|\mathbf{J}\mathbf{u} - \mathbf{g}\|_2^2 \leq \sigma^2, \tag{16}$$

where an error tolerance σ^2 is included to account for noisy data.

However, the reconstructed concentration of fluorophore, $\mathbf{u}$, must be positive, which can be imposed using a nonnegativity constraint that leads to the constrained optimization problem

$$\min_{\mathbf{u}} R(\mathbf{u}) \quad \text{such that} \quad \|\mathbf{J}\mathbf{u} - \mathbf{g}\|_2^2 \leq \sigma^2, \quad \mathbf{u} \geq 0. \tag{17}$$

The problems posed in Eqs. (16) and (17) can be solved by using conventional constrained optimization algorithms,[12,13] yet, for convex functionals, methods based on the Bregman iteration have been shown to be computationally efficient and easier to implement. In addition, they are faster and more stable than other approximations like continuation methods.[16]

II.B.4. The Bregman iteration: Data constraint

The Bregman iteration technique is based on the Bregman distance that generalizes the concept of metric associating a distance to a convex functional E not necessarily differentiable. The Bregman distance between two images $\mathbf{u}$ and $\mathbf{w}$ is defined as

$$D_E(\mathbf{u}, \mathbf{w}) = E(\mathbf{u}) - E(\mathbf{w}) - \langle \mathbf{s}, \mathbf{u} - \mathbf{w} \rangle, \tag{18}$$

where $\mathbf{s}$ is the subgradient of E at $\mathbf{w}$, a subderivative that generalizes the concept of derivative to nondifferentiable convex functionals, and that is equal to the gradient when this exists; and $\langle \cdot, \cdot \rangle$ is the scalar product. Although it is not a metric, as it does not satisfy symmetry nor the triangle inequality, it provides a generalized nonnegative measure of distance. For instance, for the case of the 2-norm functional $E(\mathbf{u}) = \|\mathbf{u}\|_2^2$ the Bregman distance reduces to the Euclidean distance $D_E(\mathbf{u}, \mathbf{w}) = \|\mathbf{u} - \mathbf{w}\|_2^2$. It can be shown that it also generalizes the Mahalanobis distance, the Itakura-Saito distance, and the Kullback-Leibler divergence.[34] For further explanations we refer to Refs. 14 and 16.

In our case of interest, let $E(\mathbf{u}) = R(\mathbf{u})$ be the total variation functional, then the constrained optimization problem [Eq. (16)] in terms of the Bregman distance has the equivalent formulation

$$\begin{aligned} \mathbf{u}^{k+1} &= \min_{\mathbf{u}} D_R(\mathbf{u}, \mathbf{u}^k) + \frac{\lambda}{2}\|\mathbf{J}\mathbf{u} - \mathbf{g}\|_2^2 \\ \mathbf{s}^{k+1} &= \mathbf{s}^k - \lambda \mathbf{J}^T(\mathbf{J}\mathbf{u}^{k+1} - \mathbf{g}) \end{aligned} \tag{19}$$

where $\mathbf{s}^k$ is the subgradient of the total variation functional at the kth-iteration.

For linear operators, $\mathbf{J}$, the Bregman formulation given by Eq. (19) has a simplified version

$$\mathbf{u}^{k+1} = \min_{\mathbf{u}} R(\mathbf{u}) + \frac{\lambda}{2}\|\mathbf{Ju} - \mathbf{g}^k\|_2^2$$
$$\mathbf{g}^{k+1} = \mathbf{g}^k + \mathbf{g} - \mathbf{Ju}^{k+1}, \tag{20}$$

with $\mathbf{g}^0 = \mathbf{g}$ the initial data. The second line is the so-called Bregman iteration, which can be explained as the requirement that imposes the constraint by adding the error of the data misfit back into the data constraint.[14,16] Thus, the introduction of the Bregman iteration into the unconstrained formulation [Eq. (13)] makes its solution converge to the solution of the constrained problem [Eq. (16)] as the number of iterations k increases, for a sufficiently small weight parameter λ. In comparison to an equivalent unconstrained formulation, it further minimizes the solution error norm by constructing a sequence of solutions that monotonically converges to the noise-free solution, with the data constraint decreasing monotonically with the iteration number.[14,16] Moreover, for linear operators, one does not need to solve the more difficult formulation given by Eq. (19), using subgradient methods for the selection of the subgradient, but can instead compute the simpler formulation given by Eq. (20).

II.B.5. The Split Bregman method: Data and nonnegativity constraints

We aim at the solution of the constrained problem [Eq. (17)] with both a data and a nonnegativity constraint using the Split Bregman formulation. SB allows the minimization of convex nondifferentiable functionals in an efficient and easy manner. Thus, the SB method leads to a solution update for which L2-functionals and the nonnegativity functional are decoupled and solved separately. To achieve this decoupling, a new variable $\mathbf{v}$ that will bear the nonnegativity condition and will allow for the splitting is introduced. Let $\Pi(\mathbf{v} \geq 0)$ be the nonnegativity condition that restricts the solution space to only positive values. Following the formalism in Ref. 16, the constrained problem given by Eq. (17) is rewritten as

$$\min_{\mathbf{u},\mathbf{v}} \Pi(\mathbf{v} \geq 0) + R(\mathbf{u}) \quad \text{such that}$$
$$\mathbf{v} = \mathbf{u} \quad \text{and} \quad \mathbf{Ju} = \mathbf{g}, \tag{21}$$

whose equivalent unconstrained formulation is

$$\min_{\mathbf{u},\mathbf{v}} \Pi(\mathbf{v} \geq 0) + R(\mathbf{u}) + \frac{\lambda}{2}\|\mathbf{Ju} - \mathbf{g}\|_2^2 + \frac{\alpha}{2}\|\mathbf{v} - \mathbf{u}\|_2^2, \tag{22}$$

Defining now a Bregman distance associated to a functional $E(\mathbf{u},\mathbf{v}) = \Pi(\mathbf{v}) + R(\mathbf{u})$, and proceeding as in Eq. (19) yet for two variables $\mathbf{u}$ and $\mathbf{v}$, the constrained problem given by Eq. (21) becomes

$$(\mathbf{u}^{k+1}, \mathbf{v}^{k+1}) = \min_{\mathbf{u},\mathbf{v}} D_E(\mathbf{u}, \mathbf{u}^k, \mathbf{v}, \mathbf{v}^k) + \frac{\lambda}{2}\|\mathbf{Ju} - \mathbf{g}\|_2^2 + \frac{\alpha}{2}\|\mathbf{v} - \mathbf{u}\|_2^2$$
$$\mathbf{s}_{\mathbf{u}}^{k+1} = \mathbf{s}_{\mathbf{u}}^k - \lambda\mathbf{J}^T(\mathbf{Ju}^{k+1} - \mathbf{g}) - \alpha(\mathbf{u}^{k+1} - \mathbf{v}^{k+1})$$
$$\mathbf{s}_{\mathbf{v}}^{k+1} = \mathbf{s}_{\mathbf{v}}^k - \alpha(\mathbf{v}^{k+1} - \mathbf{u}^{k+1}). \tag{23}$$

Using the Bregman iteration, Eq. (23) has the equivalent formulation

$$(\mathbf{u}^{k+1}, \mathbf{v}^{k+1}) = \min_{\mathbf{u},\mathbf{v}} \Pi(\mathbf{v} \geq 0) + R(\mathbf{u}) + \frac{\lambda}{2}\|\mathbf{Ju} - \mathbf{g}^k\|_2^2 + \frac{\alpha}{2}\|\mathbf{v} - \mathbf{u} - \mathbf{b}_{\mathbf{v}}^k\|_2^2$$
$$\mathbf{g}^{k+1} = \mathbf{g}^k + \mathbf{g} - \mathbf{Ju}^{k+1}$$
$$\mathbf{b}_{\mathbf{v}}^{k+1} = \mathbf{b}_{\mathbf{v}}^k + \mathbf{u}^{k+1} - \mathbf{v}^{k+1}. \tag{24}$$

with $\mathbf{g}^0 = \mathbf{g}$, $\mathbf{u}^0 = 0$, $\mathbf{v}^0 = 0$. As the variables $\mathbf{u}^{k+1}$ and $\mathbf{v}^{k+1}$ are not coupled, they can be split and solved separately

$$\mathbf{u}^{k+1} = \min_{\mathbf{u}} R(\mathbf{u}) + \frac{\lambda}{2}\|\mathbf{Ju} - \mathbf{g}^k\|_2^2 + \frac{\alpha}{2}\|\mathbf{v}^k - \mathbf{u} - \mathbf{b}_{\mathbf{v}}^k\|_2^2 \tag{25}$$

$$\mathbf{v}^{k+1} = \min_{\mathbf{v}} \Pi(\mathbf{v} \geq 0) + \frac{\alpha}{2}\|\mathbf{v} - \mathbf{u}^{k+1} - \mathbf{b}_{\mathbf{v}}^k\|_2^2. \tag{26}$$

Then, $\mathbf{u}$ in Eq. (25) is solved using conventional unconstrained optimization methods. The elements of $\mathbf{v}^{k+1}$ in Eq. (26) that are not coupled between each other are solved independently using a shrinkage formula[26]

$$\mathbf{v}^{k+1} = \max(\mathbf{u}^{k+1} + \mathbf{b}_{\mathbf{v}}^k, 0) \tag{27}$$

This is easily obtained by minimization of a quadratic functional, $\frac{\alpha}{2}\|\mathbf{v} - \mathbf{u}^{k+1} - \mathbf{b}_{\mathbf{v}}^k\|_2^2$, whose minimizer is $\mathbf{v} = \mathbf{u}^{k+1} + \mathbf{b}_{\mathbf{v}}^k$, when $\mathbf{u}^{k+1} + \mathbf{b}_{\mathbf{v}}^k > 0$, and $\mathbf{v} = 0$, when $\mathbf{u}^{k+1} + \mathbf{b}_{\mathbf{v}}^k < 0$.

II.B.6. Algorithms implementation

We compared the following algorithms.

a: GN) The unconstrained problem given by Eq. (13) solved using a GN step [Eq. (15)]

$$\mathbf{u}^{k+1} = -\left(R''(\mathbf{u}^k) + \lambda\mathbf{J}^T\mathbf{J}\right)^{-1}\left(R'(\mathbf{u}^k) + \lambda\mathbf{J}^T(\mathbf{Ju}^k - \mathbf{g})\right). \tag{28}$$

b: GN-P0) The unconstrained problem given by Eq. (13) solved using a GN step and projecting to zero the negative part of the solution at each iteration (GN-P0)

$$\mathbf{u}^{k+1} = -\left(R''(\mathbf{u}^k) + \lambda\mathbf{J}^T\mathbf{J}\right)^{-1}\left(R'(\mathbf{u}^k) + \lambda\mathbf{J}^T(\mathbf{Ju}^k - \mathbf{g})\right)$$
$$\mathbf{u}^{k+1} = 0 \quad \text{for} \quad \mathbf{u}^{k+1} < 0. \tag{29}$$

c: SB) The Split Bregman solution [Eqs. (24 26)] that iteratively converges to the solution of the constrained problem given by Eq. (17) with both data and nonnegativity constraints

$$\mathbf{u}^{k+1} = \left(\lambda\mathbf{J}^T\mathbf{J} + R''(\mathbf{u}^k) + \alpha I\right)^{-1}$$
$$\left(-R'(\mathbf{u}^k) + \lambda\mathbf{J}^T(\mathbf{Ju}^k - \mathbf{g}^k) - \alpha(\mathbf{v}^k - \mathbf{u}^k - \mathbf{b}_{\mathbf{v}}^k)\right)$$
$$\mathbf{v}^{k+1} = \max(\mathbf{u}^{k+1} + \mathbf{b}_{\mathbf{v}}^k, 0)$$
$$\mathbf{g}^{k+1} = \mathbf{g}^k + \mathbf{g} - \mathbf{Ju}^{k+1}$$
$$\mathbf{b}_{\mathbf{v}}^{k+1} = \mathbf{b}_{\mathbf{v}}^k + \mathbf{u}^{k+1} - \mathbf{v}^{k+1}, \tag{30}$$

where $\mathbf{u}^{k+1}$ corresponds to a GN minimization step of Eq. (25).

Procedures to compute the first and second derivatives of the total variation functional, $R'(\mathbf{u})$ and $R''(\mathbf{u})$, can be found in Ref. 20.

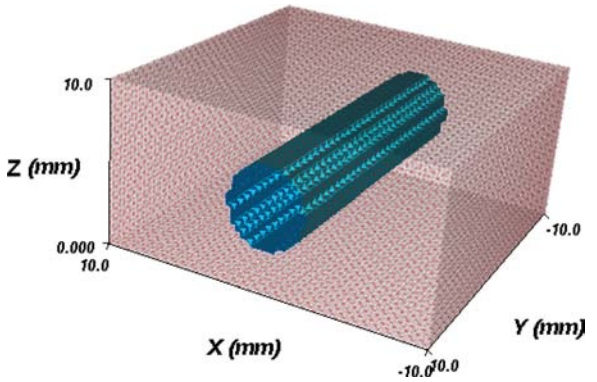


FIG. 1. Finite element model corresponding to the physical slab geometry phantom with a cylindrical region filled with fluorophore.

II.C. Method comparison

Reconstruction methods GN, GN-P0, and SB were compared using simulated and experimental data.

II.C.1. fDOT data acquisition and data sets

The fDOT system is based on a noncontact parallel plate configuration.[35,36] The sample is illuminated with a 675 nm constant-wave laser beam, focusing at desired points (source locations) using two mirrors moved by galvanometers, and the transmitted light is recorded with a CCD camera. Fluorescence and excitation images are separately recorded by using 10-nm bandwidth filters centered at 720 nm and 675 nm for fluorescent and excitation light, respectively. For each source, a distribution of detectors is selected over the CCD sensor field of view at the desired points on the sample surface. All the components of the set-up are placed inside a light-shielded box. The acquisition process is controlled by in-house developed software hosted in a PC workstation.

Experimental and simulated data corresponded to a 10 mm thick slab-geometry phantom with a cylindrical fluorophore target (Fig. 1). An experimental solid phantom was built with polyester resin, titanium oxide powder and India ink[37] ($\mu_a = 0.01$ mm^{-1}, $\mu_s' = 0.8$ mm^{-1}). A cylindrical hole of 5 mm diameter was drilled and filled with a matching fluid made of intralipid and India ink[38] mixed with Alexa fluor 750.

Computer-simulated data were generated using a fine finite element mesh (145000 nodes) (Fig. 1). A coarser mesh (55000 nodes) resampled onto a uniform mesh of $20 \times 20 \times 20$ voxels was used for the reconstruction. A 5 % additive white Gaussian noise was added to the simulated data.

The acquisition protocol used a configuration of 9 by 9 sources and detectors located at the lower and upper planes, respectively, covering a surface of 12 by 12 mm^2.

II.C.2. Criteria for the comparison of methods

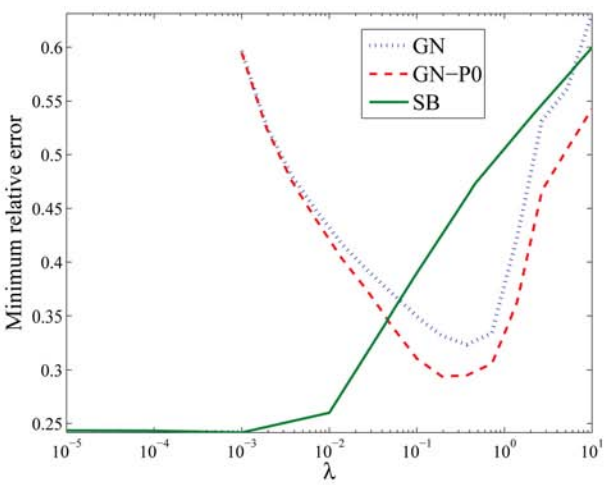
Previous to the comparison between methods, two parameters needed to be selected for all methods, the data fidelity parameter λ and the number of iterations k; in addition, for SB, there is another parameter α that weights the nonnegativity constraint. For all the methods, we used the parameter set that provided the best solution, in terms of the relative solution error norm for simulated data, and in terms of the signal-to-noise ratio for experimental data. The relative solution error norm at each iteration k was computed as[12]

$$\mathrm{err}^k = \frac{\|\mathbf{u}^k - \mathbf{u}^k_{\mathrm{true}}\|_2}{\|\mathbf{u}^k_{\mathrm{true}}\|_2}, \tag{31}$$

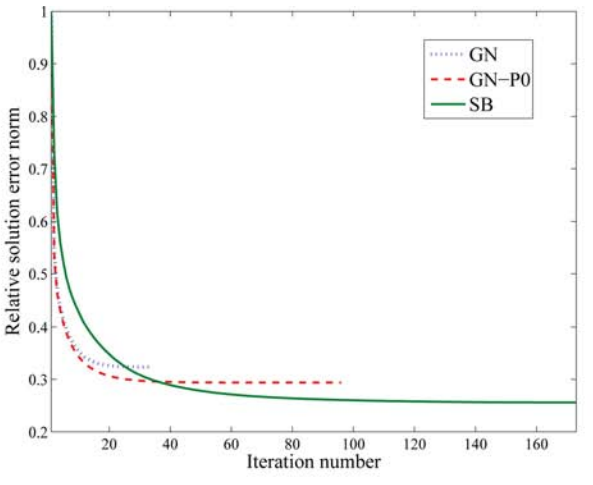
where $\mathbf{u}_{\mathrm{true}}$ is the projection on the reconstruction mesh of the target solution.

The signal-to-noise ratio was calculated as

$$\mathrm{SNR} = 10\log_{10}\frac{\|\mathbf{u}(\mathbf{r}_{\mathrm{target}})\|^2}{\|\mathbf{u}(\mathbf{r}_{\mathrm{bknd}})\|^2}, \tag{32}$$



(a)Minimum relative solution error norm versus data fidelity parameter λ



(b)Relative solution error norm versus iteration number

FIG. 2. Comparison of methods in terms of (a) minimum solution error norm for a range of the data fidelity parameter λ, and (b) relative solution error versus iteration number, for the parameter λ that minimizes the error in Fig. (a).

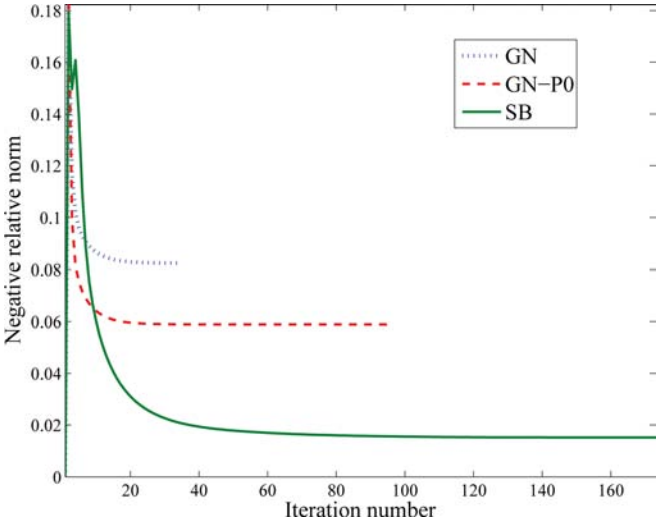


FIG. 3. Comparison of methods in terms of the negative relative part of the solution, $\|\mathbf{u}(\mathbf{u}<0)\|_2/\|\mathbf{u}\|_2$ (33), versus the iteration number, for simulated phantom data (same results as Fig. 2).

where $\mathbf{u}_{\text{true}}$ and $\mathbf{u}(\mathbf{r}_{\text{bknd}})$ are the solutions at target and background locations. The target location was approximated by selecting a small region centered at the expected location of the target, and the background location was defined by the first and last slices on the z-axis and the area surrounding the cylinder in Fig. 1.

Methods were compared in terms of minimum solution error, signal-to-noise ratio, convergence, the negative part of the solution, and image quality. Convergence was analyzed by displaying the solution error versus the iteration number. To assess for the nonnegativity constraint a negative relative norm was defined as

$$\|\mathbf{u}(\mathbf{u}<0)\|_2/\|\mathbf{u}\|_2. \tag{33}$$

In addition, axial, coronal, and sagittal views of reconstructed images and profiles along two axes are shown. The negative part of displayed images was set to zero, as negative values of concentration have no physical meaning.

III. RESULTS

III.A. Comparison of methods

III.A.1. Computer-simulated phantom data

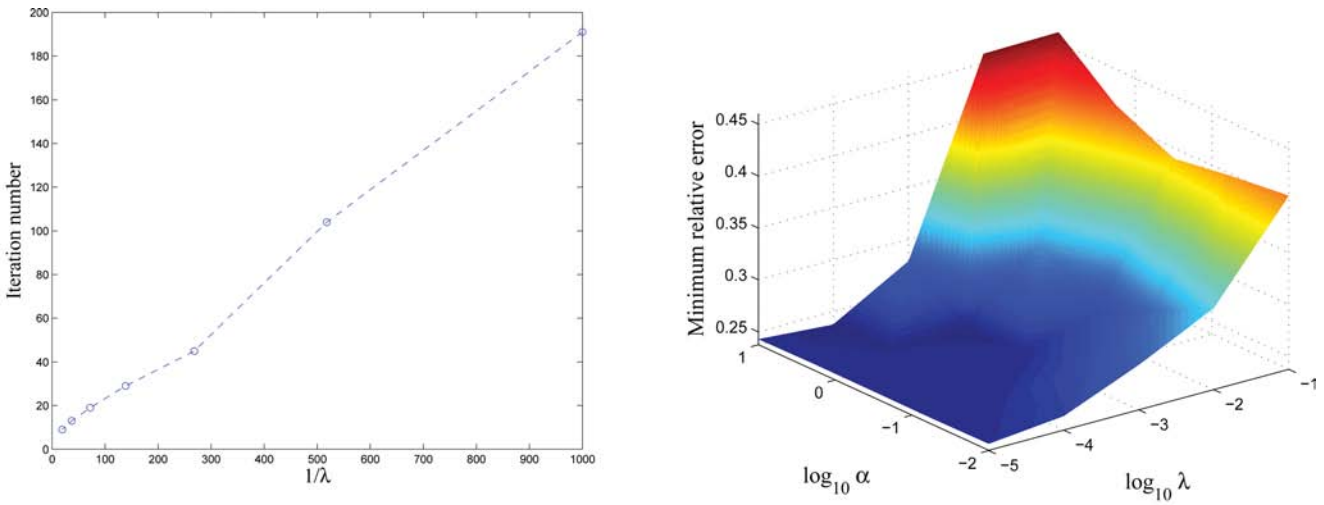
The best solution for each method was selected by choosing the data fidelity parameter λ (from the range 10^{-5} to 10) that provided the minimum solution error norm [Fig. 2(a)]. For all possible solutions, SB yielded better solution error norm than GN. GN with projection to zero yielded lower error than GN. GN converged with a fewer number of iterations than SB, whereas SB led to better solution error [Fig. 2(b)].

In terms of the negative relative norm [Eq. (33)], SB led to a solution with a noticeably lower number of negative voxels than GN-P0 (Fig. 3).

For SB, a linear relation existed between the inverse of the data fidelity parameter and the number of iterations required for convergence [Fig 4(a)]. Furthermore, SB yielded an optimum solution (with minimum error) independently of the data fidelity parameter as long as this was small enough [$\lambda \leq 10^{-3}$ in Fig. 2(a)].

Note that, while GN depends on λ, SB depends, in addition to λ, on the nonnegativity fidelity parameter α. The behavior of the error versus both parameters is plotted in Fig. 4(b).

Reconstructed images and image profiles are shown in Figs. 5 and 6. SB led to better results, with an improved



(a)Iteration number versus the inverse of the data fidelity parameter, $1/\lambda$

(b)Solution error norm versus λ and α

FIG. 4. Performance of SB for simulated phantom data. (a) Behavior of the optimum iteration number that led to optimum results (in terms of the solution error norm) versus the inverse of the data fidelity parameter λ, for a fixed value of the nonnegativity parameter ($\alpha = 10^{-1}$). (b) Solution error norm versus λ and α (same results as in Figs. 2 and 3 for SB).

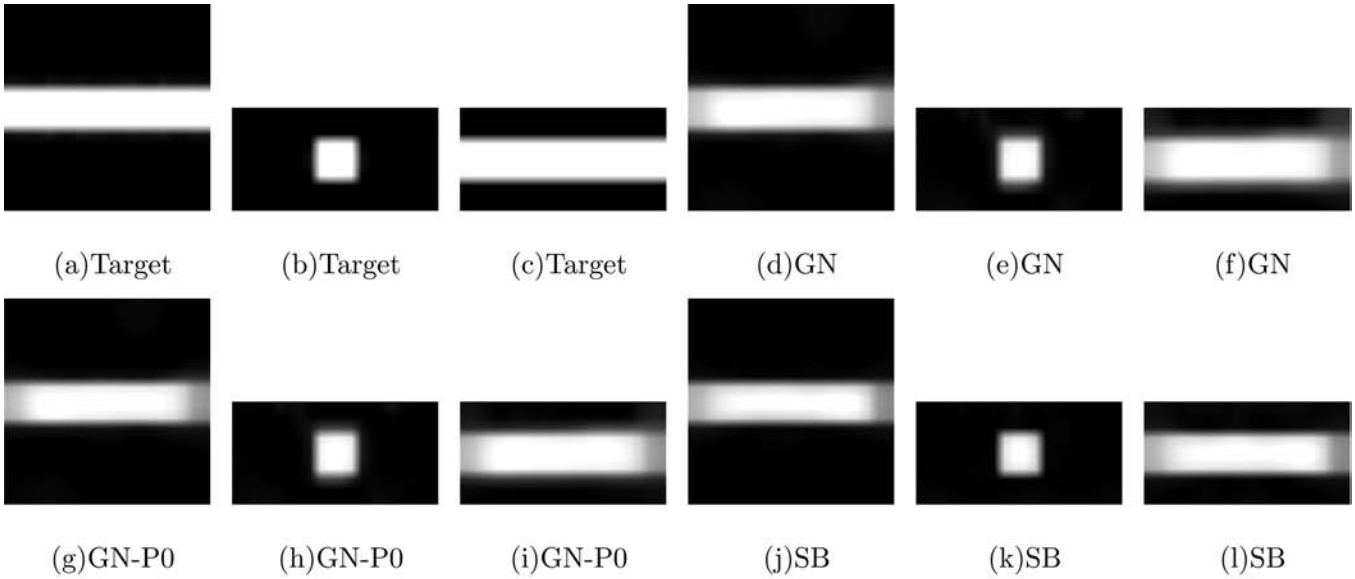


FIG. 5. Reconstructions of computer simulated phantom data with the different methods (same results as in Figs. 2 and 3). Axial, coronal, and sagittal slices (columns from left to right) for (a c) target in the reconstruction mesh, reconstructions with (d f) GN, (g i) GN P0, and (j l) SB. The negative part of images has been set to zero.

recovery of the target profile as compared to GN. For instance, the full width at tenth maximum of the image profile along the z-axis [Fig. 6(b)] was 5 mm for SB and 6.25 mm for GN. In terms of image quality, GN and GN-P0 yielded not distinguishable results, so GN-P0 was discarded for the comparison using experimental data.

III.A.2. Experimental data

In terms of SNR, the best solution for GN corresponded to $\lambda = 0.52$, selected from 15 parameters λ logarithmically spaced in the range 10^{-3} to 10; the next best solution corresponded to $\lambda = 0.32$ (the higher the data fidelity parameter λ, the better fit to the data and the less smoothing of the image). These two solutions were compared to SB.

SB led to higher SNR than GN as the number of iterations increased [Fig. 7(a)], and provided better result in terms of the negative part of the solution [Fig. 7(c)].

From the data misfit [Fig 7(b)] it is shown that GN converged to a different plateau for each value of λ. In contrast, SB continued minimizing the data misfit till it fits the noise, as a result of the Bregman iteration (explained above), and so a stopping criterion was required. A reasonable stopping criterion is a threshold on the data misfit that corresponds to the noise level. However, since an optimum threshold selection was not pursued, for the comparison we chose as

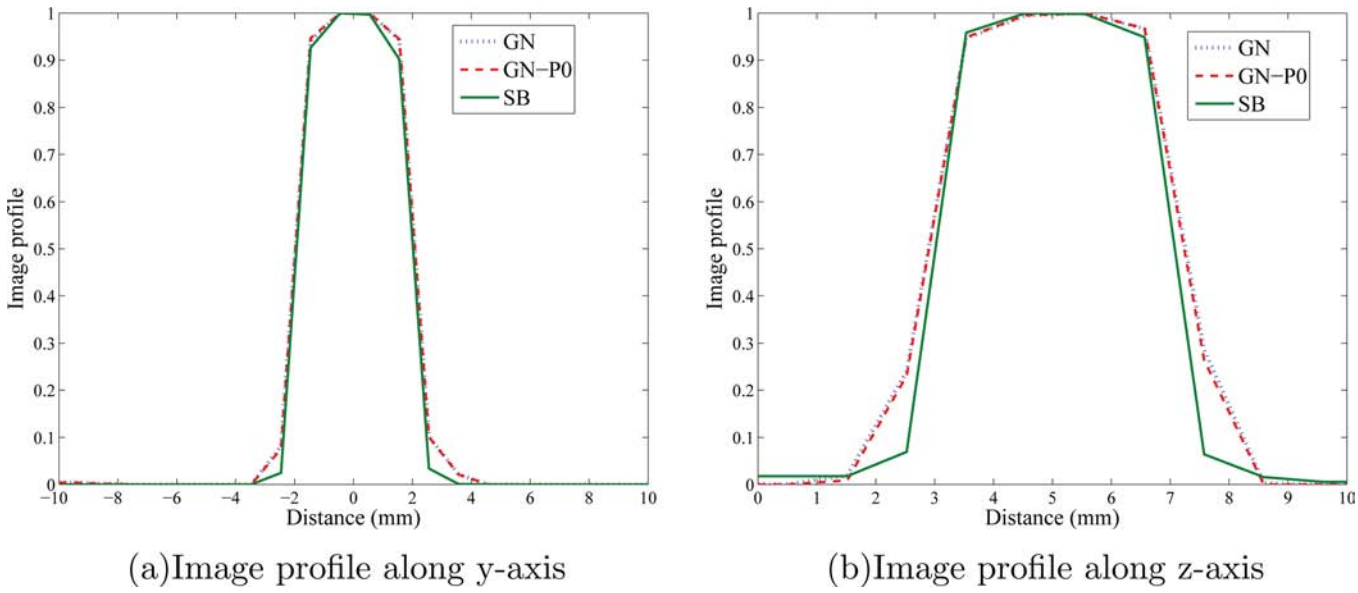


FIG. 6. Profiles along y and z axes of reconstructed images (Fig. 5) using GN, GN P0 and SB.

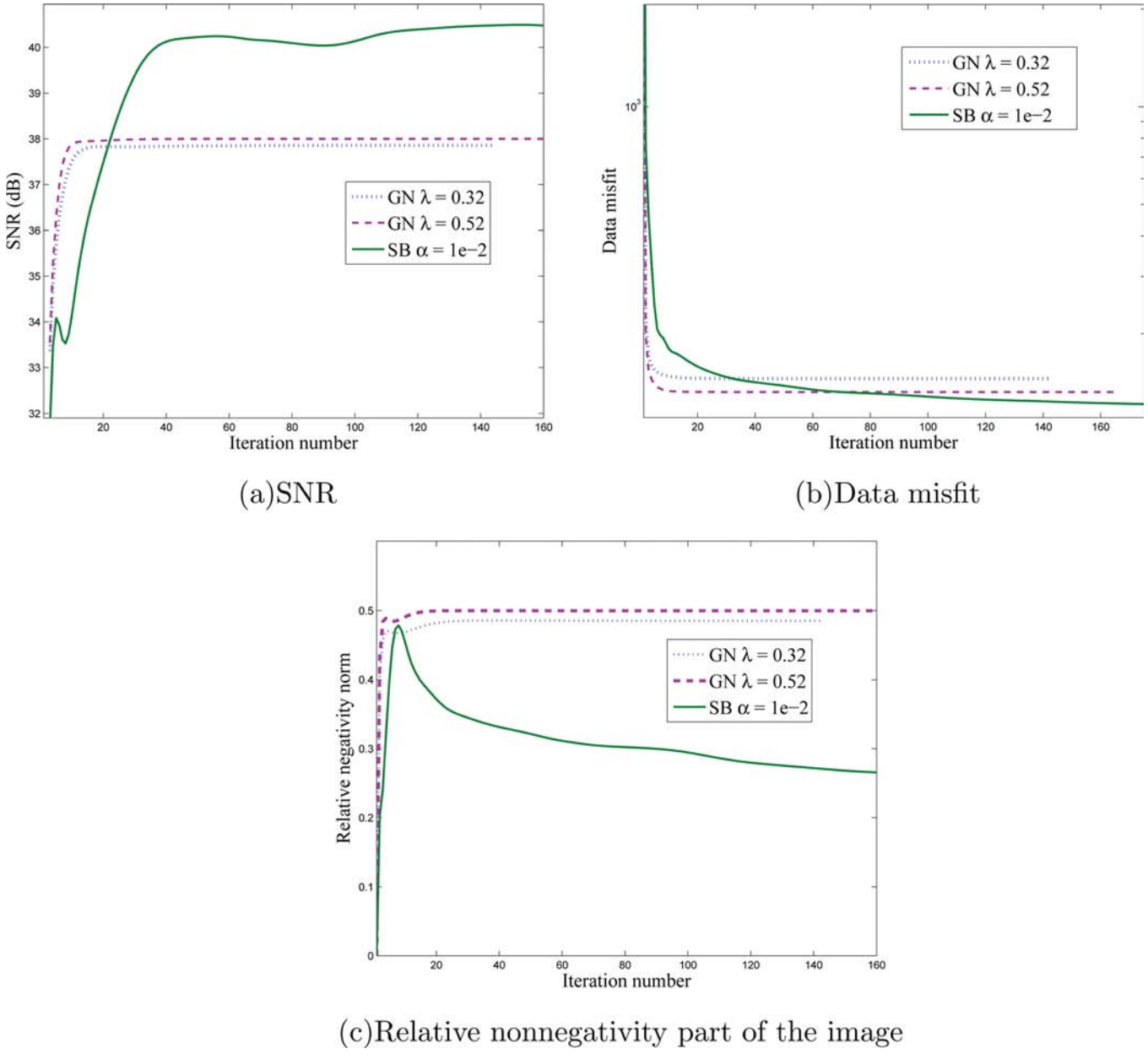


FIG. 7. Reconstruction of experimental phantom data with GN, for two data fidelity parameters (λ 0.32 and λ 0.52); and with SB, for a data fidelity parameter λ 10^{3} and a nonnegativity weighting parameter λ 10^{2}. a) Signal to noise ratio (SNR), b) data misfit $\|\mathbf{Ju}\ \mathbf{g}\|^2$, and c) the relative nonnegativity norm of the image versus the number of iterations [Eq. (33)].

threshold the data misfit given by GN at convergence. Images reconstructed with GN and SB are shown in Fig. 8.

IV. DISCUSSION

We have validated the Split Bregman algorithm as a method to reconstruct fDOT studies minimizing the total variation of the image with a nonnegativity constraint. The method has been tested on simulated and experimental phantom data. We compared SB with an equivalent unconstrained optimization approach based on a Gauss Newton step with projection to zero of the negative part of the solution after each iteration. Overall, SB was superior in terms of the solution error norm, enforcing the nonnegativity constraint, and improving image quality.

For simulated data, we found that the Bregman iteration allowed further minimization of the solution error norm. This agrees with previous reports on convergence.[14,16,17] In Ref. 14 it is explained that the Bregman iteration enforces the data constraint producing a sequence of solutions that do not stop at solving the unconstrained problem, but get closer to the true solution. Besides, we have also verified that the Bregman iteration yielded optimum results independently of the data fidelity parameter λ, for a sufficiently small parameter (which ensures large smoothing at the first iterations). In Ref. 39 it was pointed out that for a small parameter the noise is removed in the first iterations and then the fine

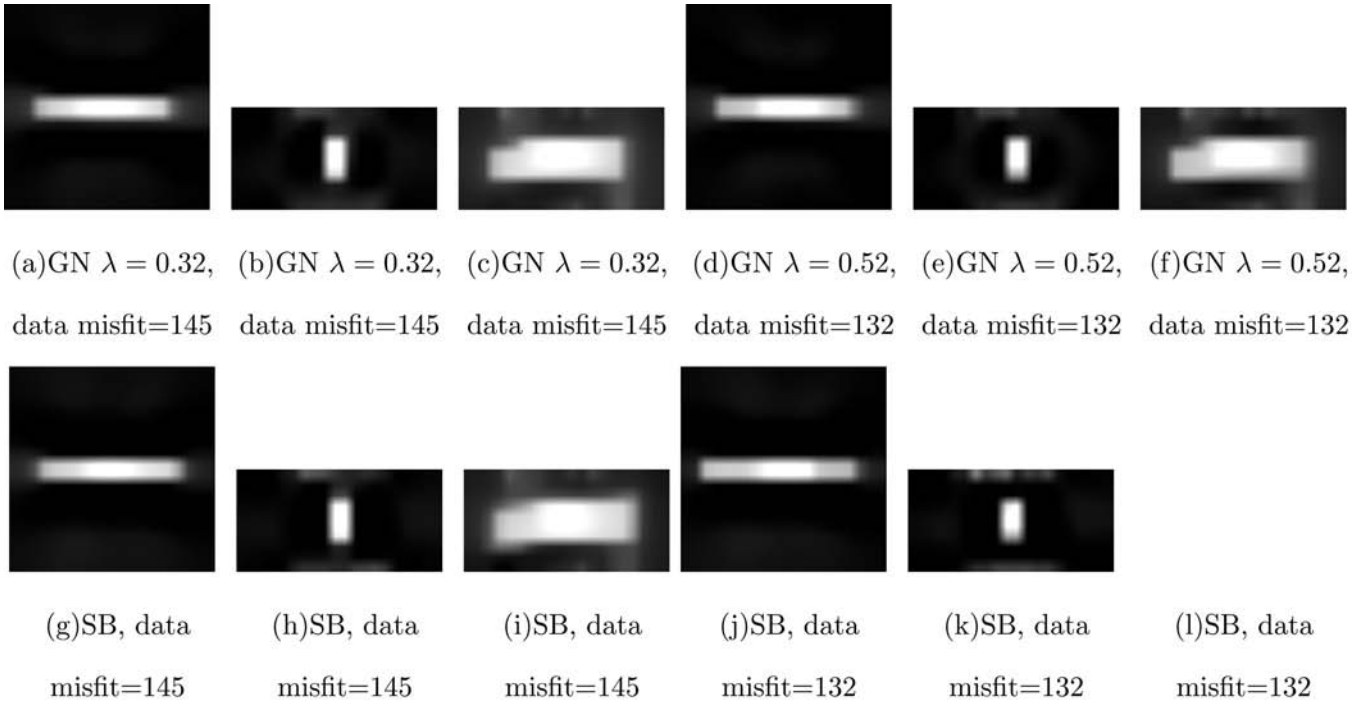


FIG. 8. Axial, coronal and sagittal slices (columns from left to right) for the reconstruction of experimental phantom data. Reconstruction with GN for data fidelity parameters (a c) λ 0.32 and (d f) λ 0.52 [Fig. 7(b)]. (g l) Reconstructions with SB (for λ 10 ^{3} and a nonnegativity weighting parameter λ 10 ^{2}) at iterations 37 and 61 that corresponded to the same data misfits than with GN in Fig. 7(b). The negative part of images has been set to zero.

scales of the image are recovered before the noise. Moreover, we verified that there is a linear relationship between the number of iterations needed to obtain optimum results and the inverse of the data fidelity parameter, which agrees with the results presented in Refs. 14 and 16.

Regarding the stopping criterion, for an equivalent unconstrained optimization problem convergence of the data misfit is reached at a different level for each λ, which requires optimum selection of λ. For SB, as long as the parameter λ is small enough, the optimum solution does not depend on the chosen λ but on the number of iterations. A stopping criterion that thresholded on the data misfit was previously used for SB.[16] Since we did not aim for an optimum stopping criterion for our experimental data, we provided two data misfit thresholds for comparison with Gauss Newton. Therefore, future implementations on experimental data would need to define an optimum stopping criterion which may differ from standard methods applied to unconstrained optimization approaches.

The main advantages of the Split Bregman method with respect to standard constrained optimization approaches are the easier implementation and computational efficiency derived from the decoupling of L1- and L2-norms. In this work we have enforced a nonnegativity constraint. Furthermore, while here we have computed first and second derivatives of the total variation functional, this can be avoided using the Split Bregman formulation as indicated in Ref. 16. Thus, adding other L1-constraints like the TV and Besov-norm is straighforward. In addition, Split Bregman could be applied to the complex linear reconstruction problem in Ref. 40 and to the nonlinear simultaneous recovery of fluorescence and background optical parameters,[41] in frequency-domain fDOT.[9]

In conclusion, we have validated the Split Bregman method for the reconstruction of simulated and experimental fDOT data, minimizing the total variation of the image with a nonnegativity constraint. Split Bregman led to an improvement in comparison to an equivalent unconstrained optimization approach, in terms of solution error and image quality.

ACKNOWLEDGMENTS

This study was supported by Ministerio de Ciencia e Innovación (FPI program, TEC 2007-64731, TEC 2008-06715, and CENIT AMIT CEN-20101014), Comunidad de Madrid and European Regional Development Fund ARTEMIS S2009/DPI-1802, and EU-FP7 project FMTXCT-201792.

[a]Author to whom correspondence should be addressed. Electronic mail: jabascal@hggm.es

[1]E. E. Graves, J. Ripoll, R. Weissleder, and V. Ntziachristos, "A submillimeter resolution fluorescence molecular imaging system for small animal imaging," Med. Phys. **30**, 901 911 (2003).
[2]V. Ntziachristos, J. Ripoll, L. V. Wang, and R. Weissleder, "Looking and listening to light: the evolution of whole body photonic imaging," Nat. Biotechnol. **23**, 313 320 (2005).
[3]S. Patwardhan, S. Bloch, S. Achilefu, and J. Culver, "Time dependent whole body fluorescence tomography of probe bio distributions in mice," Opt. Express **13**, 2564 2577 (2005).
[4]T. Lasser and V. Ntziachristos, "Optimization of 360 degrees projection fluorescence molecular tomography," Med. Image Anal. **11**, 389 399 (2007).

[5]A. Martin, J. Aguirre, A. Sarasa, D. Tsoukatou, A. Garofalakis, H. Meyer, C. Mamalaki, J. Ripoll, and A. M. Planas, "Imaging changes in lymphoid organs *in vivo* after brain ischemia with three dimensional fluo rescence molecular tomography in transgenic mice expressing green fluo rescent protein in T lymphocytes," Mol. Imaging **7**, 157 167 (2008).
[6]S. R. Arridge, M. Hiraoka, and D. T. Delpy, "A finite element approach for modelling photon transport in tissue," Med. Phys. **20**, 299 309 (1993).
[7]X. Intes, V. Ntziachristos, J. P. Culver, A. Yodh, and B. Chance, "Projection access order in algebraic reconstruction technique for diffuse optical tomography," Phys. Med. Biol. **47**, N1 N10 (2002).
[8]A. Soubret, J. Ripoll, and V. Ntziachristos, "Accuracy of fluorescent to mography in the presence of heterogeneities: study of the normalized Born ratio," IEEE Trans. Med. Imaging **24**, 1377 1386 (2005).
[9]A. Corlu, R. Choe, T. Durduran, M. A. Rosen, M. Schweiger, S. R. Arridge, M. D. Schnall, and A. G. Yodh, "Three dimensional in vivo fluo rescence diffusive optical tomography of breast cancer in humans," Opt. Express **15**, 6696 6716 (2007).
[10]D. Hyde, R. Schulz, D. Brooks, E. Miller, and V. Ntziachristos, "Performance dependence of hybrid x ray computed tomography/fluores cence molecular tomography on the optical forward problem," J. Opt. Soc. Am. A **26**, 919 923 (2009).
[11]L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," Physica D **60**, 259 268 (1992).
[12]C. R. Vogel, *Computational Methods for Inverse Problems* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 2002).
[13]J. Nocedal and S. J. Wright, *Numerical Optimization* (Springer Verlag, New York, 1999).
[14]S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, "An iterative regula rization method for total variation based image restoration," Multiscale Model. Simul. **4**, 460 489 (2005).
[15]W. Yin, S. Osher, D. Goldfarb, and J. Darbon, "Bregman iterative algo rithms for l1 minimization with applications to compressed sensing," SIAM J. Imaging Sci. **1**, 143 168 (2008).
[16]T. Goldstein and S. Osher, "The split Bregman method for L1 regularized problems," SIAM J. Imaging Sci. **2**, 323 343 (2009).
[17]E. Esser, "Applications of Lagrangian based alternating direction methods and connections to split Bregman," UCLA CAM Technical Report, 9 31 (2009).
[18]M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," Magn. Reson. Med. **58**, 1182 1195 (2007).
[19]K. D. Paulsen and H. Jiang, "Enhanced frequency domain optical image reconstruction in tissues through total variation minimization," Appl. Opt. **35**, 3447 58 (1996).
[20]A. Douiri, M. Schweiger, J. Riley, and S. R. Arridge, "Anisotropic diffu sion regularization methods for diffuse optical tomography using edge prior information," Meas. Sci. Technol. **18**, 87 95 (2007).
[21]H. Gao and H. Zhao, "Multilevel bioluminescence tomography based on radiative transfer equation Part 1: l1 regularization," Opt. Express **18**, 1854 1871 (2010).
[22]H. Gao and H. Zhao, "Multilevel bioluminescence tomography based on radiative transfer equation Part 2: Total variation and l1 data fidelity," Opt. Express **18**, 2894 2912 (2010).
[23]D. Han, J. Tian, S. Zhu, J. Feng, C. Qin, B. Zhang, and X. Yang, "A fast reconstruction algorithm for fluorescence molecular tomography with sparsity regularization," Opt. Express **18**, 8630 8648 (2010).
[24]M. Freiberger, C. Clason, and H. Scharfetter, "Total variation regulariza tion for nonlinear fluorescence tomography with an augmented lagrangian splitting approach," Appl. Opt. **49**, 3741 3747 (2010).
[25]T. Correia, J. Aguirre, A. Sisniega, J. Chamorro Servent, J. Abascal, J. J. Vaquero, M. Desco, V. Kolehmainen, and S. Arridge, "Split operator method for fluorescence diffuse optical tomography using anisotropic dif fusion regularisation with prior anatomical information," Biomed. Opt. Express **2**, 2632 2648 (2011).
[26]S. Setzer, G. Steidl, and T. Teuber, "Deblurring Poissonian images by split Bregman techniques," J. Visual Commun. Image Representation **21**, 193 199 (2010).
[27]A. B. Milstein, S. Oh, K. J. Webb, C. A. Bouman, Q. Zhang, D. A. Boas, and R. P. Millane, "Fluorescence optical diffusion tomography," Appl. Opt. **42**, 3081 3094 (2003).
[28]V. Ntziachristos, A. H. Hielscher, A. G. Yodh, and B. Chance, "Diffuse optical tomography of highly heterogeneous media," IEEE Trans. Medical Imaging **20**, 470 478 (2001).
[29]V. Ntziachristos and R. Weissleder, "Experimental three dimensional fluo rescence reconstruction of diffuse media by use of a normalized Born approximation," Opt. Lett. **26**, 893 895 (2001).
[30]R. B. Schulz, A. Ale, A. Sarantopoulos, M. Freyer, E. Soehngen, M. Zientkow ska, and V. Ntziachristos, "Hybrid system for simultaneous fluorescence and x ray computed tomography," IEEE Trans. Med. Imaging **29**, 465 473 (2010).
[31]M. Schweiger, "Application of the finite element method in infrared image reconstruction of scattering media," Ph.D. thesis, University College Lon don, London, UK, 1994.
[32]M. Schweiger, S. R. Arridge, M. Hiraoka, and D. T. Delpy, "The finite ele ment method for the propagation of light in scattering media: Boundary and source conditions," Med. Phys. **22**, 1779 1792 (1995).
[33]M. Schweiger, S. R. Arridge, and I. Nissila, "Gauss Newton method for image reconstruction in diffuse optical tomography," Phys. Med. Biol. **50**, 2365 2386 (2005).
[34]A. Banerjee, S. Merugu, and I. S. Dhillon, "Clustering with Bregman divergences," J. Mach. Learn. Res. **6**, 1705 1749 (2005).
[35]J. Ripoll, R. B. Schulz, and V. Ntziachristos, "Free space propagation of dif fuse light: Theory and experiments," Phys. Rev. Lett. **18**, 103901 (2003).
[36]J. Ripoll and V. Ntziachristos, "Imaging scattering media from a distance: theory and applications of noncontact optical tomography," Mod. Phys. Lett. B **18**, 1403 1431 (2004).
[37]D. Boas, "Diffuse Photon Probes of Structural and Dynamical Properties of Turbid Media: Theory and Biomedical Applications," Ph.D. thesis, Uni versity of Pennsylvania, Philadelphia, 1996.
[38]R. Cubeddu, A. Pifferi, P. Taroni, A. Torricelli, and G. Valentini, "A solid tissue phantom for photon migration studies," Phys. Med. Biol. **42**, 1971 1979 (1997).
[39]L. He, T. C. Chang, S. Osher, T. Fang, and P. Speier, "MR image recon struction by using the iterative refinement method and nonlinear inverse scale space methods," UCLA CAM Technical Report, 6 35 (2006).
[40]H. Gao, Y. Lin, G. Gulsen, and H. Zhao, "Fully linear reconstruction method for fluorescence yield and lifetime through inverse complex source formulation," Opt. Express **35**, 1899 1901 (2010).
[41]H. Gao, S. Osher, and H. Zhao, "Quantitative photoacoustic tomography," UCLA CAM Report, 11 28. (Mathematical Modeling in Biomedical Imaging II: Optical, Ultrasound, and Opto Acoustic Tomographies. Lec ture Notes in Mathematics), Springer Verlag, Berlin Heidelberg, **2035**, 131 158 (2012).