

RESEARCH

Open Access



An L_p ($0 \leq p \leq 1$)-norm regularized image reconstruction scheme for breast DOT with non-negative-constraint

Bingyuan Wang¹, Wenbo Wan¹, Yihan Wang¹, Wenjuan Ma², Limin Zhang¹, Jiao Li¹, Zhongxing Zhou¹, Huijuan Zhao^{1,3*} and Feng Gao^{1,3*}

*Correspondence:
huijuanzhao@tju.edu.cn;
gaofeng@tju.edu.cn
¹ College of Precision
Instruments
and Optoelectronics
Engineering, Tianjin
University, Tianjin 300072,
China
Full list of author information
is available at the end of the
article

Abstract

Background: In diffuse optical tomography (DOT), the image reconstruction is often an ill-posed inverse problem, which is even more severe for breast DOT since there are considerably increasing unknowns to reconstruct with regard to the achievable number of measurements. One common way to address this ill-posedness is to introduce various regularization methods. There has been extensive research regarding constructing and optimizing objective functions. However, although these algorithms dramatically improved reconstruction images, few of them have designed an essentially differentiable objective function whose full gradient is easy to obtain to accelerate the optimization process.

Methods: This paper introduces a new kind of non-negative prior information, designing differentiable objective functions for cases of L_1 -norm, L_p ($0 < p < 1$)-norm and L_0 -norm. Incorporating this non-negative prior information, it is easy to obtain the gradient of these differentiable objective functions, which is useful to guide the optimization process.

Results: Performance analyses are conducted using both numerical and phantom experiments. In terms of spatial resolution, quantitiveness, gray resolution and execution time, the proposed methods perform better than the conventional regularization methods without this non-negative prior information.

Conclusions: The proposed methods improves the reconstruction images using the introduced non-negative prior information. Furthermore, the non-negative constraint facilitates the gradient computation, accelerating the minimization of the objective functions.

Keywords: Diffuse optical tomography, Inverse problem, Sparsity regularization, Non-negative

Background

Diffuse optical tomography (DOT) has attracted widespread interest in recent years due to its non-invasive [1, 2] and sensitive [3] properties, which offers huge clinical potential. This biomedical imaging modality presents lower radiation risk than those methods using X-ray [3] and has extensive applications including optical mammography and functional brain imaging [1, 4]. DOT reconstructs the spatial distribution of the optical

properties such as absorption coefficients, which provide useful functional information about blood volume and oxygenation. However, the reconstructed images of DOT often suffer from the low spatial resolution and are significantly affected by measurement noise because the inverse problem of DOT is highly ill-posed [1–5]. The inverse problem of DOT is both undetermined because the available measurements are far fewer than the unknown variables to be reconstructed and ill-conditioned because of the dominance of scattering during the propagation of light in human tissues.

To alleviate this ill-posedness, researchers introduced various useful prior information to regularize the DOT inverse problem. The first introduced prior information is smoothness constrain incorporated in the well-established Tikhonov regularization method, which to some extent improved the reconstruction image. The involved optimization problem can be effectively solved because its objective function is differentiable. However, the reconstructed image of this method is often over-smoothed, due to the incorporated smoothness constraint, discouraging sharp edges in the reconstructed images. The second important prior information, used to alleviate the ill-posedness, is that the real solution is sparse. Its corresponding regularization term is based on the L_p -norm ($0 \leq p \leq 1$) instead of L_2 -norm. These sparsity regularization methods can facilitate the recovery of the sharp edges and are robust with noise [4, 5]. This is based on the clinical fact that the breast tumor usually accounts for a small part of the overall breast, and the remaining part is healthy [2, 4]. Thus, the changes of the optical coefficients, which can be caused by the regional blood flow changes or the early stage of breast cancers are expected to be localized. That is to say, the original reconstruction problem itself is sparse. Up to now, researchers have devised a variety of sparsity regularization methods using different L_p ($0 \leq p \leq 1$)-norms [6, 7]. For example, in case of L_1 -norm, Amir Beck et al. proposed a practical method named FISTA [8] using the gradient information; Figueiredo et al. proposed a method called GPSR [9] to split the L_1 -norm into two parts, from which the gradient is easy to obtain. To use L_p ($0 < p < 1$)-norm regularization, several approaches [4, 10, 11] were proposed to obtain the optimal solution with the help of majorization-minimization framework. When $p = 0$, a well-established smooth L_0 -norm based regularization method [4, 12] was designed, which had proven to be effective. The total variation methods of special sparsity regularization [13] also improved the reconstruction image by using the L_1 -norm of the difference of neighboring pixels rather than that of the vector in the reconstructed image as the constraint term.

However, image reconstruction using these algorithms is a rather lengthy process. Their objective functions are essentially non-differentiable, unlike those of the Tikhonov method, making it impossible to use their gradient information directly to minimize them. More specifically, using these algorithms requires more computational work and time seeing as the full gradient information cannot be easily obtained to guide the optimization process.

In this paper, a new kind of non-negative prior information is introduced for the first time, designing regularization methods with differentiable objective functions for L_1 -norm, L_p ($0 < p < 1$)-norm and L_0 -norm. These regularization methods not only make full use of the gradient information of the objective functions, like the Tikhonov method, but also retain the advantages of the sparsity regularizations in improving image quality. To investigate the performances of these proposed methods, the methods with the non-negative constraint and those without the non-negative constraint are compared, using numerical simulation and phantom experiments.

Methods

The non-negative prior information

This work assumes that scattering coefficients are both known and spatially constant. Utilizing the finite element formulation to discretize the photon propagation model, the reconstruction of the absorption coefficients in DOT can be simplified into a linear equation with the following form:

$$\delta\Gamma = \mathbf{J}\delta\boldsymbol{\mu}_a + \mathbf{n}, \quad (1)$$

where $\delta\Gamma \in \mathbf{R}^{N_m}$ is the measurement vector with the number of measurements N_m . $\mathbf{J} \in \mathbf{R}^{N_m \times N_n}$ is the Jacobian matrix with the number of nodes N_n in the finite element method. $\delta\boldsymbol{\mu}_a \in \mathbf{R}^{N_n}$ refers to the perturbation of absorption coefficients. $\mathbf{n} \in \mathbf{R}^{N_m}$ is the additive noise introduced unavoidably during the measurement or the error in the computation.

The common regularization methods, using different prior information to solve the above linear equation, can be summarized with an optimization problem as follows

$$\delta\boldsymbol{\mu}_a^* = \arg \min_{\delta\boldsymbol{\mu}_a} \left\{ \frac{1}{2} \|\mathbf{J}\delta\boldsymbol{\mu}_a - \delta\Gamma\|_2^2 + \lambda \|\delta\boldsymbol{\mu}_a\|_p^p \right\}, \quad (2)$$

where the first term $\frac{1}{2} \|\mathbf{J}\delta\boldsymbol{\mu}_a - \delta\Gamma\|_2^2$ is the fidelity term and the second term $\lambda \|\delta\boldsymbol{\mu}_a\|_p^p$ is the regularization term. The regularization parameter, λ , is a non-negative value balancing between the data fitting and L_p -norm penalty ($0 \leq p \leq 1$). When $p = 2$, Eq. (2) is the well-known Tikhonov method. Equation (2) represents a variety of sparsity regularization methods, when $0 \leq p \leq 1$. The fidelity term is differentiable but the regularization term is non-differentiable when $0 \leq p \leq 1$, making it impossible to effectively minimize the objective function using the full gradient information.

To optimize Eq. (2) more efficiently and more precisely, a new kind of non-negative prior information is incorporated to constrain the solution. Specially speaking, from the viewpoint of physiology, the abnormal absorption coefficients caused by the breast tumor with angiogenesis must have bigger absorption coefficients than those of the normal region because the tumor has more hemoglobin [5]. Thus the changes of the absorption coefficients compared with the normal region are non-negative; that is to say, $\delta\boldsymbol{\mu}_a \geq \mathbf{0}$ can be used as an inequality constraint. This is the non-negative prior information, which will play important role in designing differentiable objective functions for sparsity regularization methods and reconstructing images of higher quality.

L_1 -norm regularized reconstruction scheme with the non-negative constraint

When $p = 1$, the regularization term can be made differentiable by introducing the non-negative constraint. Specifically speaking, knowing $\delta\boldsymbol{\mu}_a \geq \mathbf{0}$ in advance, Eq. (2) can be simplified into the following simple problem

$$\delta\boldsymbol{\mu}_a^* = \arg \min_{\delta\boldsymbol{\mu}_a \geq \mathbf{0}} \left\{ f(\delta\boldsymbol{\mu}_a) = \frac{1}{2} \|\mathbf{J}\delta\boldsymbol{\mu}_a - \delta\Gamma\|_2^2 + \lambda \sum_{i=1}^{N_n} \delta\boldsymbol{\mu}_{ai} \right\}. \quad (3)$$

The objective function $f(\delta\boldsymbol{\mu}_a)$ in (3) is differentiable, making it possible to optimize (3) using full one-order gradient information. The gradient of the objective function is $\mathbf{g} = \mathbf{J}^T * (\mathbf{J}\delta\boldsymbol{\mu}_a - \delta\Gamma) + \lambda \mathbf{1}_{N_n}$, where $\mathbf{1}_{N_n} = [1, 1, \dots, 1]^T \in \mathbf{R}^{N_n \times 1}$. The constraint

$\delta\boldsymbol{\mu}_a \geq \mathbf{0}$ is a bound constraint, for which the gradient projection method [9] is particularly effective, further simplifying the optimization process. Before exhibiting in detail the full algorithm, the projection of any arbitrary vector $\delta\boldsymbol{\mu}_a$ onto the feasible region $[\mathbf{L}, \mathbf{U}]$ is defined as follows: the i th component is given by

$$P(\delta\boldsymbol{\mu}_a, \mathbf{L}, \mathbf{U})_i = \begin{cases} \mathbf{L}_i & \text{if } \delta\mu_{ai} < \mathbf{L}_i \\ \delta\mathbf{U}_{ai} & \text{if } \delta\mu_{ai} \in [\mathbf{L}_i, \mathbf{U}_i] \\ \mathbf{U}_i & \text{if } \delta\mu_{ai} > \mathbf{U}_i \end{cases} \quad (4)$$

The piecewise linear path $\delta\boldsymbol{\mu}_a(t)$ starting from the initial point and obtained by projecting the steepest descent direction at $\delta\boldsymbol{\mu}_a^0$ onto the feasible region is thus given by

$$\delta\boldsymbol{\mu}_a(t) = P(\delta\boldsymbol{\mu}_a^0 - t\mathbf{g}, \mathbf{L}, \mathbf{U}), \quad (5)$$

where $\mathbf{g} = \nabla f(\delta\boldsymbol{\mu}_a^0)$ and t is the stepsize. $\delta\boldsymbol{\mu}_a \geq \mathbf{0}$ is a bound constraint; that is to say, the upper bound is $\mathbf{U} = [\infty, \infty, \dots, \infty]^T \in \mathbf{R}^{N_{n \times 1}}$ and the lower bound is $\mathbf{L} = [0, 0, \dots, 0]^T \in \mathbf{R}^{N_{n \times 1}}$. This process is simple and requires less amount of computation, making it suitable for solving large-scale problems where the dimensions are very large.

The detailed gradient projection method for L_1 -norm regularized reconstruction scheme with non-negative constraint is shown below, according to the optimization theory. Hereinafter, this method will be referred to as the NL_1 (Non-Negative L_1) method for short.

Pseudo code of the NL_1 method:

Initialize : Choose starting point $\delta\boldsymbol{\mu}_a^k = \mathbf{0}$, stopping criteria $\varepsilon = 1e-4$, maximum

number of iteration $M = 5000$; Set current iterate number $k = 0$;

While ($k < M$)

Project reconstructed results onto the feasible region $[\mathbf{L}, \mathbf{U}]$: $\delta\boldsymbol{\mu}_a^k = P(\delta\boldsymbol{\mu}_a^k, \mathbf{L}, \mathbf{U})$;

Compute search direction:

$$\mathbf{d}^k = \mathbf{J}\delta\boldsymbol{\mu}_a^k - \delta\boldsymbol{\Gamma}; \mathbf{g}^k = \mathbf{J}^T * \mathbf{d}^k + \lambda \mathbf{1}_{N_n}; \mathbf{p}^k = -\mathbf{g}^k; (\mathbf{d}^k \text{ is the residual vector , } \mathbf{g}^k$$

is the gradient , \mathbf{p}^k is the search direction)

$$\text{if } \|\mathbf{d}^k\|_2^2 / \|\mathbf{d}^k - \mathbf{d}^{k-1}\|_2^2 \leq \varepsilon$$

Return $\delta\boldsymbol{\mu}_a^k$ as the final result and end iterating;

else

$$\text{Compute stepsize: } t = -\frac{\mathbf{g}^{kT} \mathbf{p}^k}{\mathbf{p}^{kT} \mathbf{J}^T \mathbf{J} \mathbf{p}^k};$$

$$\text{Update the reconstructed results: } \delta\boldsymbol{\mu}_a^k = \delta\boldsymbol{\mu}_a^k + t\mathbf{p}^k;$$

Update the current iterate number: $k = k + 1$;

End if

End while

L_p -norm regularized reconstruction scheme with the non-negative constraint

In the case of $0 < p < 1$, the gradient of the regularization term is different from that of the L_1 -norm. Taking advantage of the non-negative prior information, the proposed objective function is

$$\delta \boldsymbol{\mu}_a^* = \arg \min_{\delta \boldsymbol{\mu}_a \geq \mathbf{0}} \left\{ f(\delta \boldsymbol{\mu}_a) = \frac{1}{2} \|\mathbf{J} - \delta \boldsymbol{\Gamma}\|_2^2 + \lambda \sum_{i=1}^{N_n} \delta \boldsymbol{\mu}_{ai}^p \right\} \quad (6)$$

In this scenario, the gradient of the objective function is $\mathbf{g} = \mathbf{J}^T * (\mathbf{J} \delta \boldsymbol{\mu}_a - \delta \boldsymbol{\Gamma}) + \lambda \sum_{i=1}^{N_n} \frac{1}{\delta \boldsymbol{\mu}_{ai}^{1-p} + C}$, where C is a small positive constant that provides stability by ensuring that the zero components of $\delta \boldsymbol{\mu}_a$ do not prohibit a nonzero estimate in the next step. Like the NL_1 method algorithm, the NL_p (Non-Negative L_p) algorithm is designed by replacing the gradient term with $\mathbf{g} = \mathbf{J}^T * (\mathbf{J} \delta \boldsymbol{\mu}_a - \delta \boldsymbol{\Gamma}) + \lambda \sum_{i=1}^{N_n} \frac{1}{\delta \boldsymbol{\mu}_{ai}^{1-p} + C}$. The remaining part of this algorithm is the same as that of the NL_1 algorithm. In this paper, the values of p and C are fixed at $1/2$ and $1e-6$, respectively.

L_0 -norm regularized reconstruction scheme with the non-negative constraint

For $p = 0$, i.e., the NL_0 (Non-Negative L_0) method, the optimization problem is NP-hard, and cannot be solved efficiently. Incorporating the non-negative prior information and the approximate L_0 -norm, the proposed complete objective function is as follows

$$\delta \boldsymbol{\mu}_a^* = \arg \min_{\delta \boldsymbol{\mu}_a \geq \mathbf{0}} \left\{ f(\delta \boldsymbol{\mu}_a) = \frac{1}{2} \|\mathbf{J} - \delta \boldsymbol{\Gamma}\|_2^2 + \lambda \sum_{i=1}^{N_n} \left[1 - e^{-\frac{\delta \boldsymbol{\mu}_{ai}^2}{2\sigma^2}} \right] \right\}. \quad (7)$$

The parameter σ controls the degree of approximation. For smaller values of σ , the regularization term is more approximate to L_0 -norm and is more difficult to be minimized [12]. After ignoring constant terms, Eq. (7) can be rewritten as

$$\delta \boldsymbol{\mu}_a^* = \arg \min_{\delta \boldsymbol{\mu}_a \geq \mathbf{0}} \left\{ f(\delta \boldsymbol{\mu}_a) = \frac{1}{2} \|\mathbf{J} - \delta \boldsymbol{\Gamma}\|_2^2 + \lambda \sum_{i=1}^{N_n} \left[-e^{-\frac{\delta \boldsymbol{\mu}_{ai}^2}{2\sigma^2}} \right] \right\}. \quad (8)$$

The introduction of the non-negative prior information makes it easy to calculate the gradient as

$$\mathbf{g} = \mathbf{J}^T * (\mathbf{J} \delta \boldsymbol{\mu}_a - \delta \boldsymbol{\Gamma}) + \lambda \sum_{i=1}^{N_n} \frac{\delta \boldsymbol{\mu}_{ai}}{\sigma^2} e^{-\frac{\delta \boldsymbol{\mu}_{ai}^2}{2\sigma^2}}, \quad (9)$$

The whole algorithm for the NL_0 (Non-Negative L_0) is presented as follows using the pseudo code:

Pseudo code of the NL₀ method:

Initialize : Set current outer iterate number $j=0$ and inner iterate number $k=0$;

Choose starting point $\delta\boldsymbol{\mu}_a^k = (\mathbf{J}^T \mathbf{J} + \lambda \mathbf{I})^{-1} \mathbf{J}^T \delta\boldsymbol{\Gamma}$ as the minimum L₂-norm solution of $\delta\boldsymbol{\Gamma} = \mathbf{J} \delta\boldsymbol{\mu}_a$. Set stopping criteria to $\varepsilon_1 = 1e-4$; and

$\varepsilon_2 = 1e-4$; Set the maximum number of outer iteration to

$Iter_outer = 5000$ and the inner iteration to $Iter_inner = 5$. Select the

starting value for the decreasing sequence of σ as

$$\sigma_1 = 4 * \max(|\delta\boldsymbol{\mu}_a^0|)$$

While_outer $j < Iter_outer$

While_inner ($k < Iter_inner$)

Project reconstructed results onto the feasible region $[\mathbf{L}, \mathbf{U}]$

$$\delta\boldsymbol{\mu}_a^k = \mathbf{P}(\delta\boldsymbol{\mu}_a^k, \mathbf{L}, \mathbf{U});$$

Compute search direction:

$$\mathbf{d}^k = \mathbf{J} \delta\boldsymbol{\mu}_a^k - \delta\boldsymbol{\Gamma}; \mathbf{g}^k = \mathbf{J}^T * \mathbf{d}^k + \lambda \sum_{i=1}^{N_s} \frac{\delta\boldsymbol{\mu}_{ai}}{\sigma_j^2} e^{-\frac{\delta\boldsymbol{\mu}_{ai}^2}{2\sigma_j^2}}; \mathbf{p}^k = -\mathbf{g}^k; ($$

\mathbf{d}^k is the residual vector , \mathbf{g}^k is the gradient , \mathbf{p}^k is the search direction)

if $\|\mathbf{d}^k\|_2^2 / \|\mathbf{d}^k - \mathbf{d}^{k-1}\|_2^2 \leq \varepsilon_1$

Return the $\delta\boldsymbol{\mu}_a^k$ to the outer loop;

$k=0$;

Break;

else

$$\text{Compute stepsize: } t = -\frac{(\mathbf{g}^k)^T \mathbf{p}^k}{(\mathbf{p}^k)^T \mathbf{J}^T \mathbf{J} \mathbf{p}^k};$$

Update the reconstructed results: $\delta\boldsymbol{\mu}_a^k = \delta\boldsymbol{\mu}_a^k + t\mathbf{p}^k$;

Update the current iterate number: $k = k + 1$;

End if

End while_inner

if $|\frac{\delta\boldsymbol{\mu}_a^j - \delta\boldsymbol{\mu}_a^{j-1}}{\delta\boldsymbol{\mu}_a^j}| < \varepsilon_2$

Return the $\delta\boldsymbol{\mu}_a^j$ as the final result and break.

else

$$\sigma_{j+1} = 0.6 * \sigma_j;$$

End while_outer

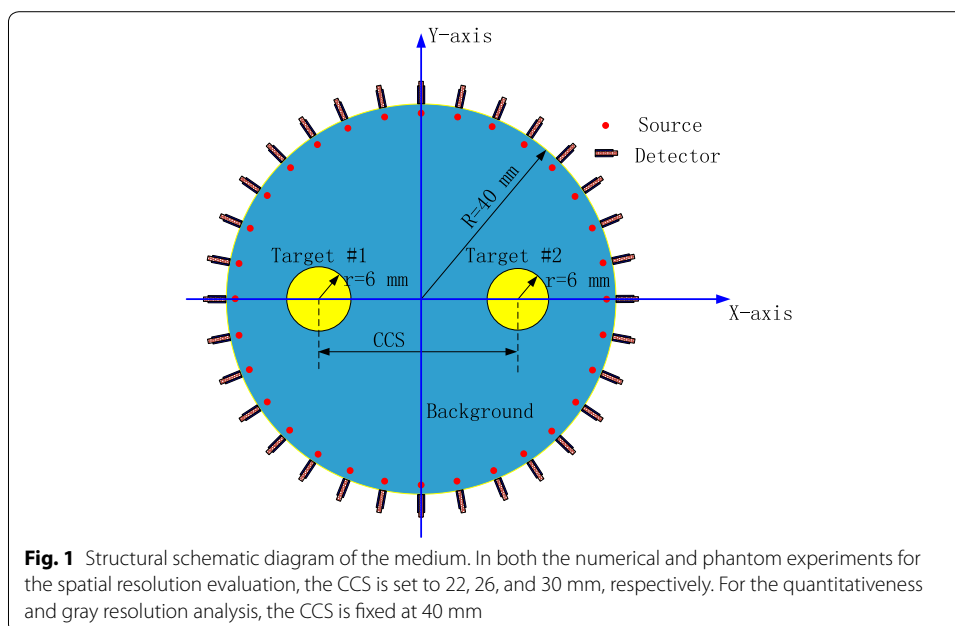
Numerical and phantom experiments

To evaluate the performances of the proposed regularization methods, several numerical and phantom experiments are conducted. Their reconstructed results are evaluated and then compared with those of the conventional regularization methods based on the same norm.

Conditions

As illustrated in Fig. 1, a 2-D circular breast-sized turbid medium with the radius of $R = 40$ mm is used for the investigations. Two circular tumor-emulating regions of 6-mm radius (Target #1 and Target #2) are embedded into the background medium symmetrical to the Y-axis, and have higher absorption coefficients. Depending on the purposes of the experiments, the center-to-center separation (CCS) between the two targets can be varied. In both the numerical and phantom experiments for the spatial resolution evaluation, the targets are located at $(x = \pm 11$ mm, $y = 0$ mm), $(x = \pm 13$ mm, $y = 0$ mm) and $(x = \pm 15$ mm, $y = 0$ mm), i.e., $CCS = 22$, 26 and 30 mm, respectively. In the experiments for the quantitiveness and gray resolution assessments, the targets are fixed at $(x, y) = (x = \pm 20$ mm, $y = 0$ mm), corresponding to $CCS = 40$ mm.

The optical properties of healthy and diseased breast tissues have been extensively investigated using both in vitro and in vivo methods [14]. Most investigations have measured the absorption and (reduced) scattering coefficients of healthy breast tissues in the range of 0.002 – 0.008 mm^{-1} and 0.6 – 1.3 mm^{-1} , respectively, in the near-infrared wavelength range. The in vivo measurements reported by Tromberg et al. indicate that some tumors exhibit 1.25 to threefold higher absorption than normal breast tissue [15]. Overall, most evidence suggests that the ratio between the absorption for healthy and diseased tissues are of the order of a factor of 2 [14]. In this paper, to simulate the clinical cases, the absorption and (reduced) scattering coefficients of the background medium



are set to 0.004 and 1 mm⁻¹, respectively. For the target regions, the absorption coefficients vary with the purposes of the experiments while the scattering coefficients are the same as that of the background: for the spatial resolution, the absorption coefficient of the targets are set as 0.008 mm⁻¹; for the quantitiveness, both the two targets have the same absorption coefficients of 0.005, 0.006 mm, and 0.008 mm⁻¹, mimicking the tumors with contrasts of 1.25, 1.5 and 2.0, respectively. For the grayscale resolution, the two targets have different absorption coefficients, paired as (0.0072, 0.0088 mm⁻¹), (0.0064, 0.0096 mm⁻¹) and (0.0056, 0.0104 mm⁻¹), corresponding to a grayscale difference of 10, 20 and 30%. It is worth noting that the same experimental setup, including the geometry and optical coefficients, are also used in the phantom experiments, facilitating the mutual collation between the numerical and phantom experiments.

Thirty-two source positions (red regions) evenly located one mean transport length inside the boundary are illuminated by the continuous-wave source one by one. For each illumination, except the 15 detectors nearest to the source position, the remaining 17 detectors evenly located on the surface of the medium are used to measure the intensities of the reemitted light, leading to a total of 32 × 17 measurements.

Four metrics are adopted to objectively select the optimal regularization parameters based on the reconstruction image quality [16]. First, the root mean square error (RMSE) is defined to measure the difference between the reconstructed and the true values

$$RMSE = \sqrt{\frac{\|\mu_a(\mathbf{r}) - \mu_a^{\text{tr}}(\mathbf{r})\|_2^2}{\|\mu_a^{\text{tr}}(\mathbf{r})\|_2^2}}, \quad (10)$$

where $\mu_a(\mathbf{r})$ and $\mu_a^{\text{tr}}(\mathbf{r})$ are the reconstructed and true absorption images at position \mathbf{r} , respectively. Second, the area ratio (AR) is defined to measures the ratio of the reconstructed target area to the true one

$$AR = \frac{|A^{\text{tg}}|}{|A^{\text{tr}}|}, \quad (11)$$

where A^{tr} and A^{tg} denotes the true and reconstructed target regions, respectively, with the latter defined as the voxels with their values higher than one half of the maximum of the reconstructed values, and $|\cdot|$ is the operator for area calculation. Third, the contrast-to-noise ratio (CNR) is defined to assess how the targets can be distinguished from the background

$$CNR = \frac{\text{Mean}_{\mathbf{r} \in A^{\text{tg}}}[\mu_a(\mathbf{r})] - \text{Mean}_{\mathbf{r} \in A^{\text{bg}}}[\mu_a(\mathbf{r})]}{\sqrt{w \text{Var}_{\mathbf{r} \in A^{\text{tg}}}[\mu_a(\mathbf{r})] + (1 - w) \text{Var}_{\mathbf{r} \in A^{\text{bg}}}[\mu_a(\mathbf{r})]}}, \quad (12)$$

where $w = |A^{\text{tg}}|/|A^{\text{tg}}| + |A^{\text{bg}}|$ with A^{bg} representing the background region, ‘Mean’ and ‘Var’ operators for mean and variance calculations, respectively. Finally, to globally assess the reconstruction image quality, a total error (TE) is defined as follows:

$$TE = e^{RMSE + \text{absolute}(AR)} / CNR, \quad (13)$$

When optimizing the regularization parameter λ , TE is calculated as a function of the regularization parameter, λ . The λ producing the smallest TE is selected because it provides the best balance across the above 3 metrics.

For enforcing evaluation of the method performances, three further metrics, namely SR , AC and GR , are defined to quantitate the spatial resolution, quantitateness and grayscale discerning ability of the reconstructed images, respectively [17]

$$SR = \frac{\mu_a(x)|_{x=0} - \min_x[\mu_a(x)]}{\max_x[\mu_a(x)] - \min_x[\mu_a(x)]}, \quad (14)$$

$$AC = \frac{\max_x[\mu_a(x)]}{\min_x[\mu_a(x)]}, \quad (15)$$

$$GR = \frac{\max_{x \geq 0}[\mu_a(x)] - \max_{x \leq 0}[\mu_a(x)]}{\max_{x \geq 0}[\mu_a(x)] + \max_{x \leq 0}[\mu_a(x)]}, \quad (16)$$

where $\mu_a(x)$ is the profile of the absorption image along the X-axis, i.e., $\mu_a(x) = \mu_a(\mathbf{r})|_{\mathbf{r}=(x,0)}$.

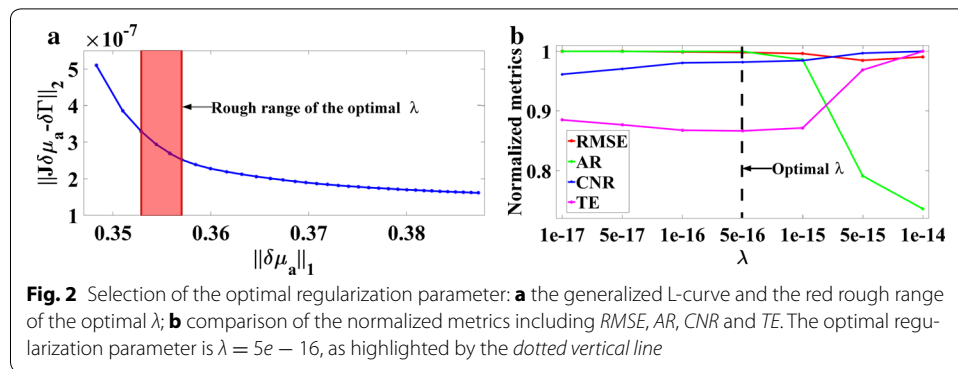
Generally speaking, better image reconstruction can be identified by: smaller RMSE, TE, and SR values; a larger CNR value; an AR value closer to 1; and AC and GR values closer to the true values.

Numerical experiments

To more closely simulate the real case, we assume that the main noises in the measurements are Gaussian noise [4–6]. Gaussian of different levels are added to the pure simulated measurements. The signal-to-noise ratio (SNR) of the detecting location with the weakest light is set to $SNR_{\min} = 20$ and 30 dB, respectively. The best signal-to-noise ratio for a number of photons, N , reaching the detector in a given time interval is $SNR = \sqrt{N}$ [18]. So, the $SNRs$ for the other detecting locations are set according to $SNR = SNR_{\min} \sqrt{I/I_{\min}}$, with I and I_{\min} being the intensities of the current detecting locations and the one with the weakest light, respectively. Different meshes are adopted in the forward and inverse problems to avoid tricky problems. Specifically, a fine mesh with 12,290 nodes and 24,098 triangle elements and a coarse mesh with 4526 nodes and 8762 triangle elements are used for the forward and inverse calculations, respectively. To comprehensively analyze and compare the methods, we repeat each experiment 10 times and plot the average value and 95% confidence interval of the different metrics.

Spatial resolution

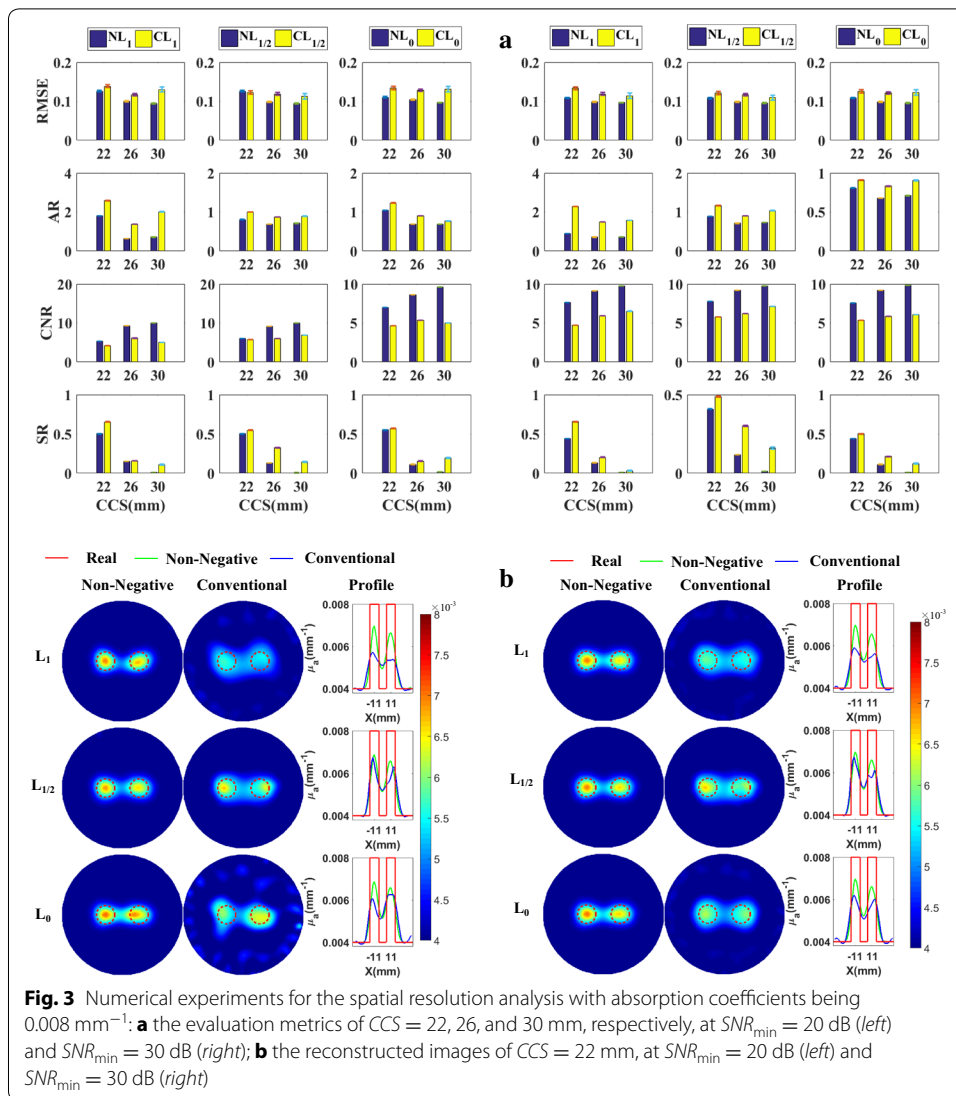
Before demonstrating the reconstructed results, we first illustrate an example about the process to select the optimal regularization parameter λ for the NL_p ($p = 1/2$) method in the spatial resolution experiments with a fixed $CCS = 22$ mm. This process consists of two steps. The first step is to find a rough range of optimal λ utilizing the generalized L-curve, shown in Fig. 2a. It traces, for a specific pair of \mathbf{J} and $\delta\mathbf{I}$, the optimal tradeoff in the space covered by the L_2 -norm of residual and the L_p -norm of the regularized solution. Like the L-curve of the Tikhonov regularization method, by locating the corner of the generalized L-curve, an approximation to the optimal regularization parameter can



be obtained [19]. The corner is located in the range of $[1e-17 \ 1e-14]$ (The red region). In the second step, more regularization parameters in this rough range are investigated and their normalized metrics are plotted in Fig. 2b. From Fig. 2b, it is clear that *AR* decreases gradually as λ increases when $\lambda \leq 1e - 15$, but decreases dramatically when $\lambda \geq 1e - 15$. This indicates that too large a λ excessively localizes the targets so that the reconstructed targets are much smaller than the real ones. The *RMSE* decreases gradually in the whole range investigated, indicating that the difference between the reconstructed results and the true ones becomes smaller. The *CNR* increases as λ increases, meaning that the reconstructed targets are easier to be distinguished from the background. Most importantly, the *TE*, valuing the reconstruction error in total, first decreases slowly and then increases rapidly. In summary, $\lambda = 5e - 16$ corresponding to the smallest *TE*, is selected as the optimal regularization parameter. It is worth noting that the reconstruction is insensitive to the change of λ when $\lambda \leq 1e - 15$. So, the method is robust with the regularization parameter λ smaller than $1e-15$. For all the following experiments, this method to select the optimal regularization parameter is used.

Figure 3a shows the column diagram of the metrics of the reconstructed images using different regularization methods. The first, second, and third columns present the metrics of the methods based on the L_1 -norm, $L_{1/2}$ -norm, and L_0 -norm respectively. Different rows contain different metrics labeled on the left side. The left and right parts contain the metrics for $SNR_{\min} = 20$ dB and $SNR_{\min} = 30$ dB, respectively. All the following figures presenting the metrics are likewise presented. From Fig. 3a, we can observe that, excepting *AR*, all metrics of the non-negative methods are better than or comparable to those of the conventional methods based on the same norm. From the confidence intervals of the metrics, we can see that the non-negative methods are more robust with the noise.

Furthermore, to more clearly and directly demonstrate the results, we plot the reconstructed images for the cases of $CCS = 22$ mm in Fig. 3b. These are the most difficult cases to separate the two nearest targets and the best cases to highlight the differences of spatial resolution abilities between different methods. The first and second columns present the images reconstructed using the non-negative methods and the conventional methods without the non-negative constraint, respectively. The third column contains the profiles of the reconstructed absorption coefficients along the X-axis. The first, second, and third rows contain the results corresponding to the L_1 , $L_{1/2}$ and

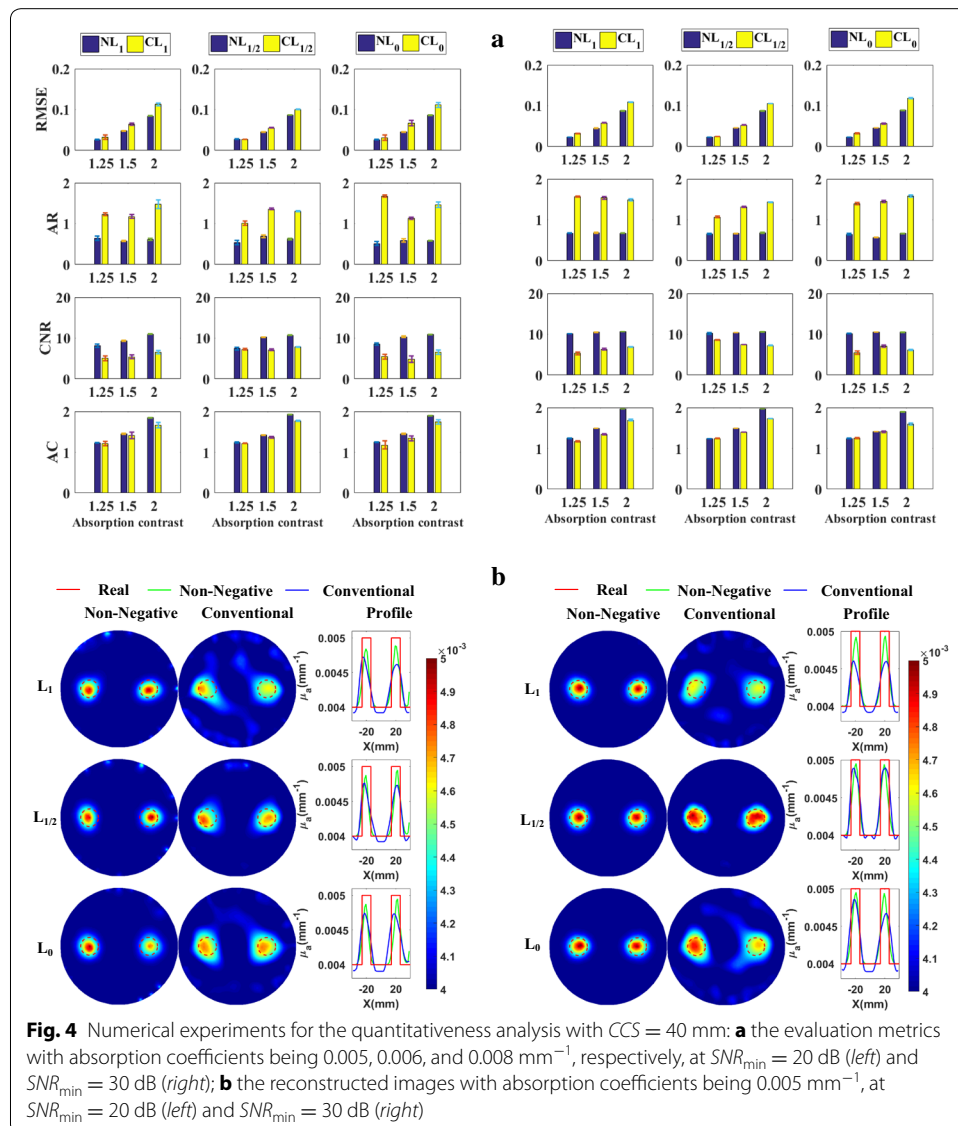


L_0 -norm, respectively. To facilitate the comparison between the methods with and without the non-negative prior information, all the images are shown using the same color-map. The profiles for the same norm-based regularization methods are plotted in the same subfigures. The left and right parts contain the metrics for $\text{SNR}_{\min} = 20 \text{ dB}$ and $\text{SNR}_{\min} = 30 \text{ dB}$, respectively. All the following figures presenting the images are likewise presented. From Fig. 3b, it is clear that the targets reconstructed using the proposed methods have smaller sizes and are better separated than those of the conventional methods based on the same norm. At the same time, the results of the non-negative regularization methods have less artifacts than those of the conventional methods. This validates the bigger CNR of the non-negative regularization methods compared to those of the conventional methods. This is because the proposed methods better suppress the influences of the noise and of the difference in the meshes used in the forward and inverse problems. Figure 3a, b jointly demonstrate that the non-negative regularization methods have better spatial resolution while retaining better RMSE and CNR .

Quantitativeness

To focus on studying the quantitativeness of the methods, the *CCS* is fixed at 40 mm, which is far enough for the methods to separate the two targets. The metrics of the reconstructed images using different regularization methods are shown in Fig. 4a. Excepting the *AR*, all metrics of the non-negative methods are better than those of the conventional methods based on the same norm.

Furthermore, to more clearly and directly demonstrate the results, we plot the reconstructed images for the cases of $AC = 1.25$ in Fig. 4b. These are the most difficult cases to distinguish the targets from the background, due to the low contrast, and the best cases to highlight the difference of quantitativeness between different methods. From Fig. 4b, we observe that the images of the non-negative regularization methods have less artifacts and are better distinguished from the background than those of the conventional methods based on the same norm. Figure 4a, b jointly demonstrate that the

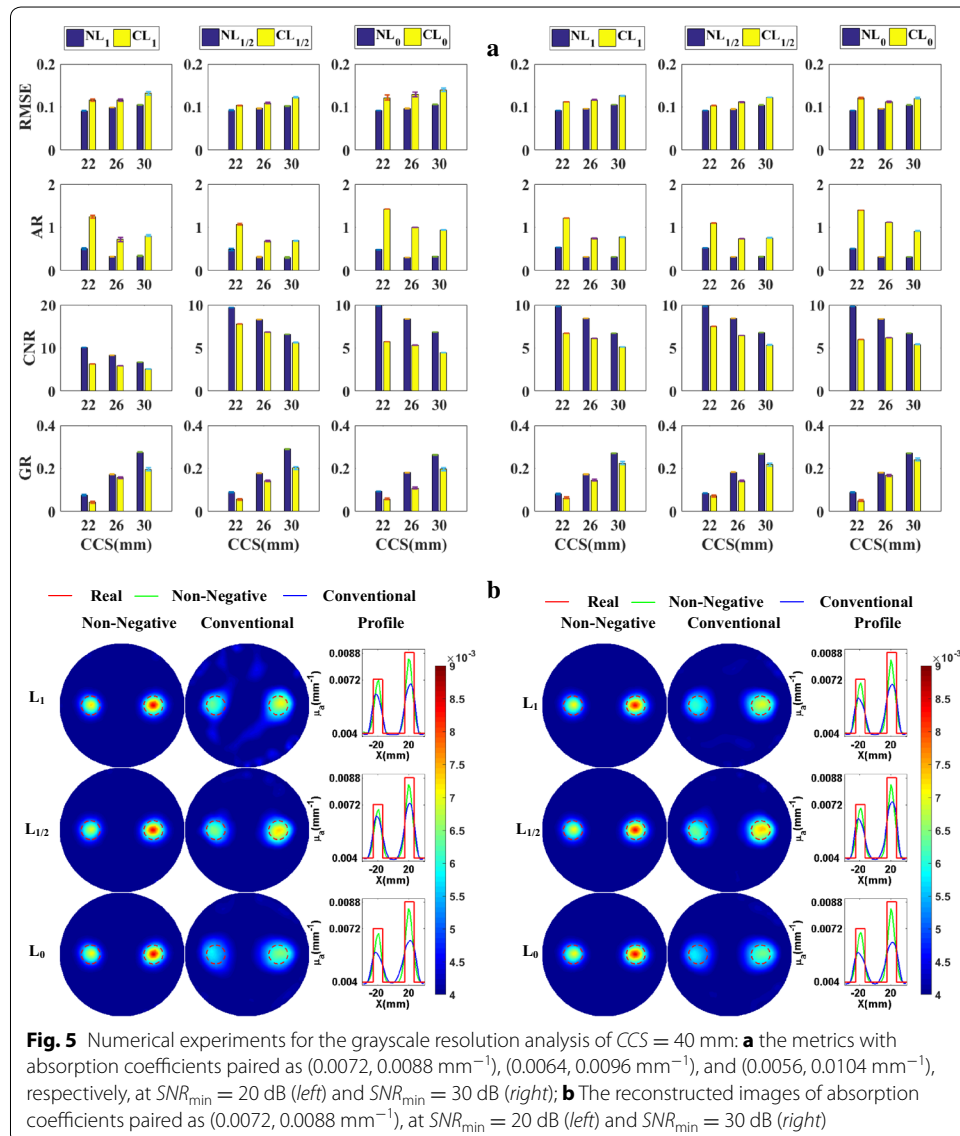


non-negative regularization methods have better quantitiveness abilities while retaining better *RMSE* and *CNR*.

Grayscale resolution

The metrics of the reconstructed images using different regularization methods are plotted in Fig. 5a. Figure 5a shows that all the metrics, except the *AR*, of the reconstructed images of the non-negative methods are better than those of their conventional methods based on the same norm.

Furthermore, to more clearly and directly demonstrate the results, we plot the reconstructed images for the cases with a *GR* being 10% in Fig. 5b. These are the most difficult cases to reconstruct the small grayscale difference, and the best cases to highlight the difference of grayscale resolution between different methods. It is easy to find that targets reconstructed using the non-negative methods have less artifacts and closer grayscale



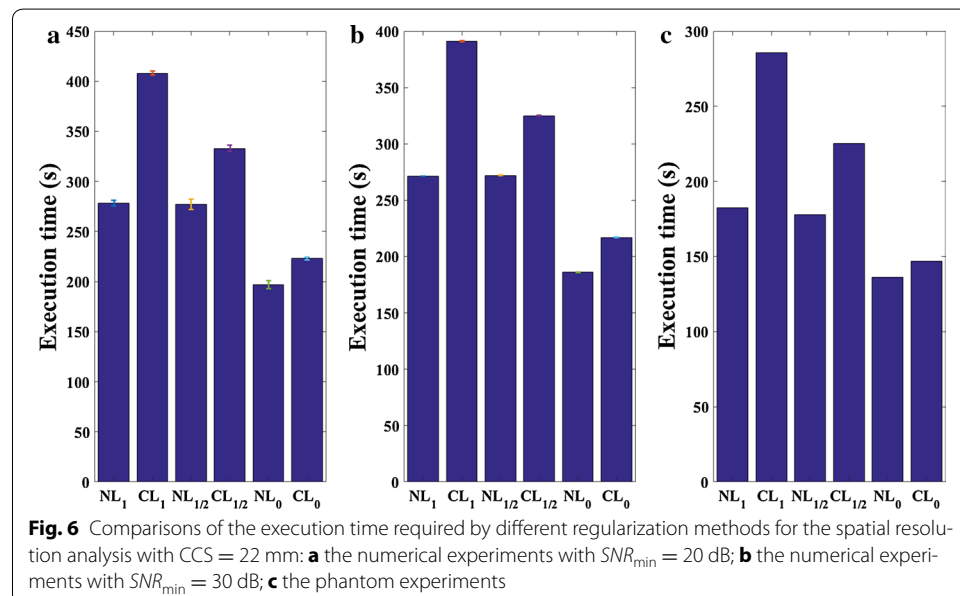
difference to the real values. Figure 5a, b jointly demonstrate that the non-negative regularization methods have better grayscale resolution ability while retaining better *RMSE* and *CNR*.

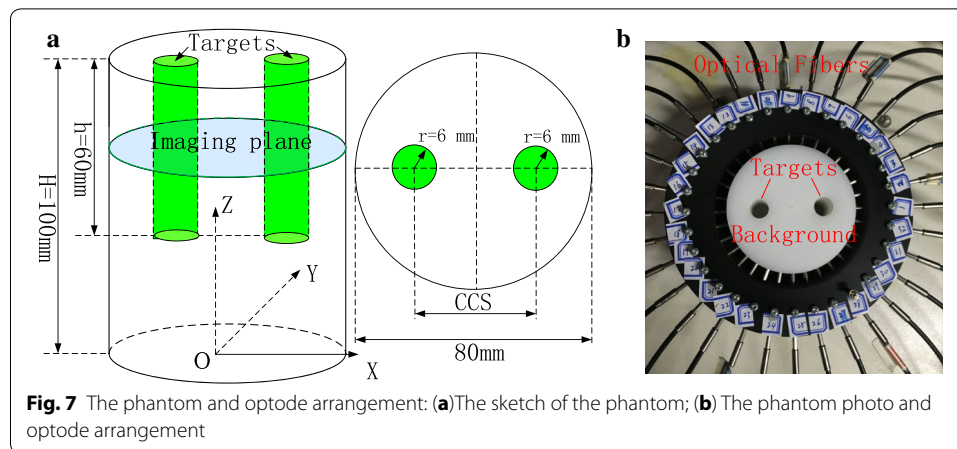
Execution time

To compare the speed of the proposed methods, all the involved methods are investigated in the same numerical experiments. During the previous numerical experiments, analyzing the spatial resolution for the case of $CCS = 20$ mm, the execution time for all methods are recorded. In these experiments, the stopping criteria are the same, as shown in the algorithms of NL_1 , $NL_{1/2}$ and NL_0 . The same computer, Interl(R) Core(TM) i7-4790 CPU @3.60 GHz, is used in all these experiments. The execution time averaged over 10 repeated experiments and the 95% confidence interval are listed in Fig. 6a, b. From Fig. 6a, b, we can see that the non-negative regularization methods not only obtain images with better quality but also require shorter time than the conventional methods based on the same norm.

Phantom experiment

To further investigate the methods proposed in this paper, a continuous wave system [20] is used to conduct phantom experiments. A cylindrical polyoxymethylene phantom with the same cross section as that of the numerical experiments is adopted. The cross section of this phantom is discretized into a mesh with 4526 nodes and 8762 triangle elements, identical to that of the numerical scenario. The detailed structure of this phantom is shown in Fig. 7a. The phantom photo and optode arrangement are shown in Fig. 7b. Via the time domain system [21], the absorption and (reduced) scattering coefficients of this phantom at a wavelength of 675 nm are measured as 0.004 and 1 mm^{-1} , identical to those of the background regions in numerical simulations. As described in the *Condition* section, the same experiments as those of the numerical experiments are conducted for the following phantom experiments. It is worth noting that we just conduct these





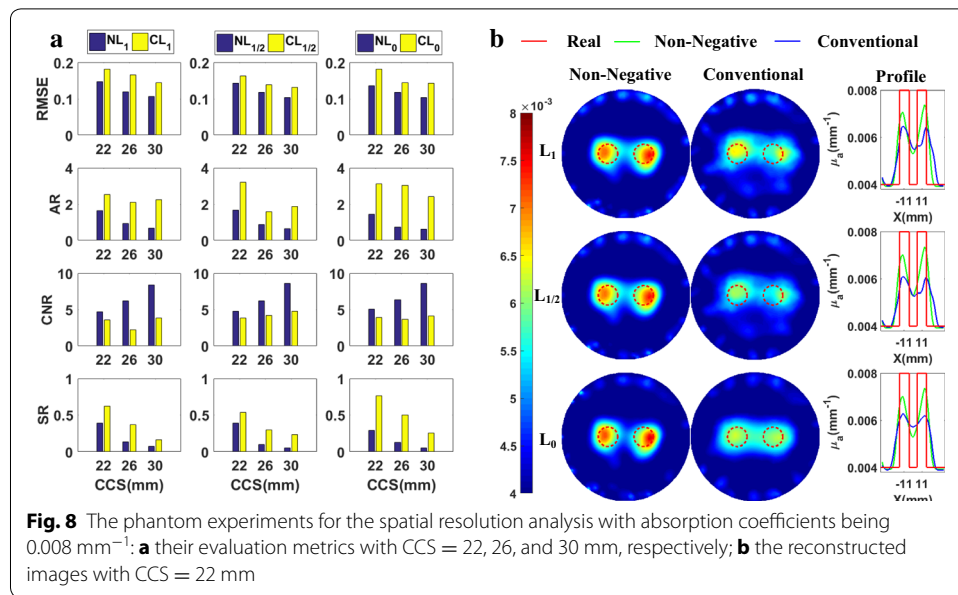
phantom experiments once time and do not consider the experiments with different levels of noises.

Along the boundary of cylindrical phantom, 32 optodes with a coaxial structure of source and detector are placed equidistantly from one another (as shown in Fig. 7b). A steady state laser operating at a wavelength of 675 nm stimulates the phantom boundary through a 32:1 mechanical optical switch. Meanwhile, the reemitted light on the phantom boundary is collected and sent to 4 photomultiplier tubes (PMT) through four 8:1 mechanical optical switches. Then the single electron responses from the PMTs are detected and accumulated by the field programmable gate array (FPGA). The total number accumulated in 250 ms is used as the intensity of the light on the detecting locations on the phantom boundary. It is worth noting that the light of the locations near the sources is so strong that it possibly induces the stack effect, making the measurement imprecise. Therefore, only the remaining opposite 17 measurements farthest away from the sources are used. A total of $32 \times 17 = 544$ measurements, equal to that in the numerical experiments, are then used to reconstruct the spatial distribution of absorption coefficients.

All phantom experiments are conducted according to the following process. First, the holes embedded in the phantom are filled with the mixed solution having absorption and (reduced) scattering coefficients of 0.004 and 1 mm^{-1} , identical to those of the healthy breast tissue. One of the optical fibers works as a light source and the remaining opposite 17 optical fibers are used to measure the light intensity reemitted from the surface of the phantom. Then, the remaining optical fibers work as the light source one by one, and the corresponding light intensities are measured. The mixed solution in the target holes is then replaced with the mixed solution with larger absorption coefficients, mimicking the breast tumors of different absorption contrasts. Repeating the measurement process, another 544 measurements are obtained and then are utilized in the reconstruction process.

Spatial resolution analysis

The metrics of the reconstructed images are plotted in Fig. 8a. This figure shows that, excepting the AR, the metrics of the reconstructed images of the non-negative



regularization methods are better than those of the conventional methods without the non-negative constraint based on the same norm.

Furthermore, to more clearly and directly demonstrate the results, we plot the reconstructed images for the cases of $CCS = 22 \text{ mm}$ in Fig. 8b. These are the most difficult cases to separate the two nearest targets and the best cases to highlight the differences of spatial resolution abilities between different methods. Figure 8b tells us that the images of the non-negative regularization methods have less artifacts and the two targets are better separated than those of the conventional methods. Figure 8a, b jointly demonstrate that the non-negative regularization methods have better spatial resolution ability while retaining better *RMSE* and *CNR*.

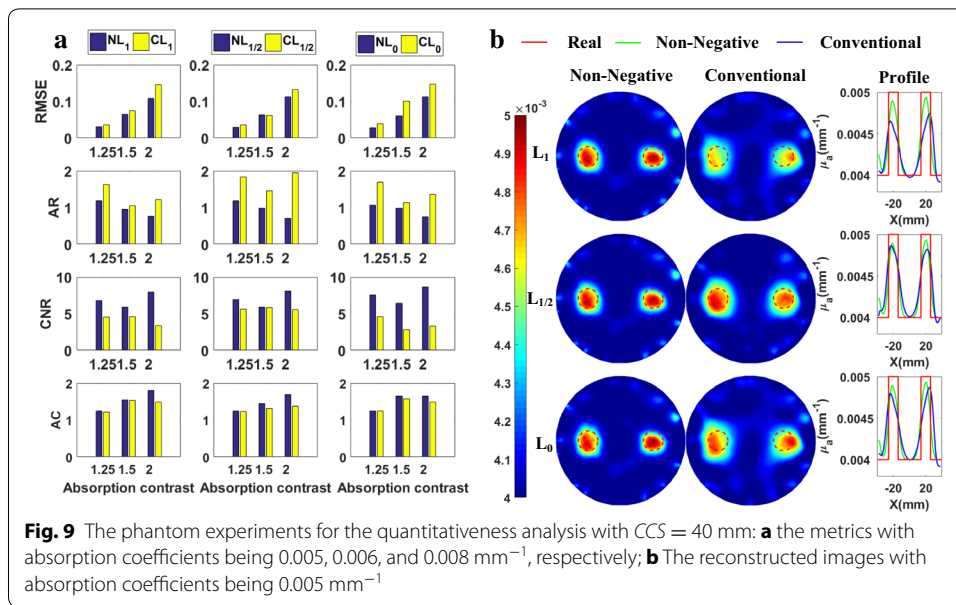
Quantitativeness

We plot the metrics of the images reconstructed using the different methods in Fig. 9a. It is clear that, the metrics, except the *AR*, of the reconstructed images of non-negative methods are better than those of the conventional methods based on the same norm.

Furthermore, to more clearly and directly demonstrate the results, we plot the reconstructed images for the cases of $AC = 1.25$ in Fig. 9b. These are the most difficult cases to distinguish the targets from the background, due to the low contrast, and the best cases to highlight the difference of quantitativeness between different methods. From Fig. 9b, we observe that the images of the non-negative regularization methods have less artifacts and are better distinguished from the background than those of the conventional methods. Figure 9a, b jointly demonstrate that the non-negative regularization methods have better quantitativeness ability while retaining better *RMSE* and *CNR*.

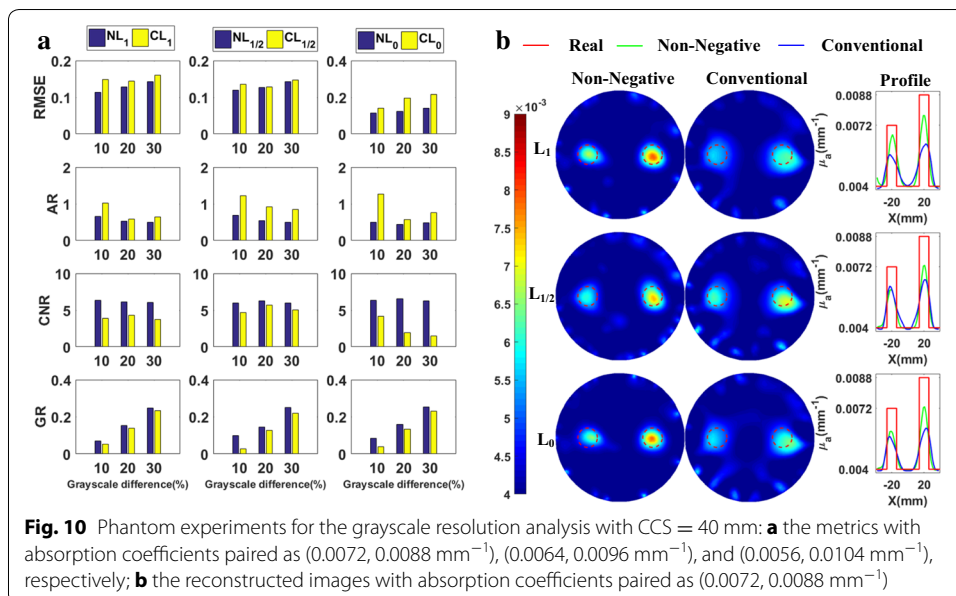
Grayscale resolution

Figure 10a contains the metrics of the images reconstructed using the different regularization methods. From this figure, it is clear that the metrics, except the *AR*, of the



non-negative regularization methods are better than those of the conventional regularization methods based on the same norm.

Furthermore, to more clearly and directly demonstrate the results, we present the reconstructed images for the cases of a grayscale difference of 10% in Fig. 10b. These are the most difficult cases to reconstruct the small grayscale difference, and the best cases to highlight the difference of grayscale resolution between different methods. Figure 10b shows that the targets reconstructed using the non-negative methods have less artifacts and closer grayscale difference to the real values than those of the conventional methods. Figure 10a, b jointly demonstrate that the non-negative regularization methods have better grayscale resolution abilities while retaining better *RMSE* and *CNR*.



Execution time

The same stopping criteria and computer as those of the numerical simulation experiments are also adopted. Execution time of all the regularization methods for the phantom experiments analyzing spatial resolution for $CCS = 22$ mm are recorded and presented in Fig. 6c. It is obvious that the non-negative regularization methods require less time to obtain the optical solution than those of the conventional regularization methods based on the same norm.

Discussion and conclusion

Though the investigations described in this paper are implemented on the continuous wave breast diffuse optical tomography, they can easily be extended to diffuse fluorescence tomography, because they share the same sparse and non-negative prior information.

The gradient projection methods described above can be replaced with a logarithmic barrier method. For the L_1 -norm, the involved unconstrained optimization problem is as follows:

$$\delta\boldsymbol{\mu}_a^* = \operatorname{argmin} \left\{ P(\delta\boldsymbol{\mu}_a; \xi) = \frac{1}{2} \|\mathbf{J}\delta\boldsymbol{\mu}_a - \delta\boldsymbol{\Gamma}\|_2^2 + \lambda \sum_{i=1}^{N_n} \delta\mu_{ai} - \frac{1}{\xi} \log(\delta\boldsymbol{\mu}_a) \right\} \quad (17)$$

where ξ is referred to as the barrier parameter. The gradient of the objective function is $\mathbf{J}^T(\mathbf{J}\delta\boldsymbol{\mu}_a - \delta\boldsymbol{\Gamma}) + \lambda\mathbf{1}_{N_n} - \frac{1}{\xi} \frac{1}{\delta\boldsymbol{\mu}_a + C}$, which is easy to compute. However, it requires too much time to find an appropriate barrier parameter ξ , making it unpractical. This logarithmic barrier method also applies to the non-negative L_p -norm ($0 < p < 1$) and L_0 -norm based regularization methods, but they also require much time to find the suitable barrier parameter.

In conclusion, we propose several new regularized methods based on L_1 -norm, L_p -norm ($0 < p < 1$) and approximate L_0 -norm with non-negative prior information. Both numerical and phantom experiments demonstrate that, excepting *AR*, all metrics are improved by these proposed methods. Furthermore, the proposed methods require less time than those of the conventional methods based on the same norm.

Abbreviations

DOT: diffuse optical tomography; CCS: center-to-center separation; RMSE: root mean square error; AR: area ratio; CNR: contrast-to-noise ratio; TE: total error; SR: spatial resolution; AC: absorption contrast; GR: grayscale resolution; SNR_{\min} : the minimum signal noise ratio; PMT: photomultiplier tubes; FPGA: field programmable gate array; Non-Neg- L_1 : L_1 -norm regularized reconstruction scheme with non-negative constraint.

Authors' contributions

BW carried out the design of the algorithms and the experiments. WW and YW participated in the implementation of the experiments. WM, LZ, JL and ZZ participated in providing guidance for designing the algorithm. HZ and FG carried out the design of comparison of the different regularization methods. All authors read and approved the final manuscript.

Author details

¹ College of Precision Instruments and Optoelectronics Engineering, Tianjin University, Tianjin 300072, China. ² Cancer Institute and Hospital, Tianjin Medical University, Tianjin 300060, China. ³ Tianjin Key Laboratory of Biomedical Detecting Techniques and Instruments, Tianjin 300072, China.

Acknowledgements

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Data availability statement

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Funding

The authors acknowledge the funding supports from the National Natural Science Foundation of China (81371602, 61475115, 61475116, 81401453, 61575140, 81571723, 81671728), and Tianjin Municipal Government of China (14JCQNJC14400, 15JCZDJC31800, 15JCQNJC14500, 16JCZDJC31200).

Received: 20 September 2016 Accepted: 30 January 2017

Published online: 03 March 2017

References

1. Cao N, Nehorai A, Jacob M. Image reconstruction for diffuse optical tomography using sparsity regularization and expectation-maximization algorithm. *Opt Express*. 2007;15:13695–708.
2. Chen C, Tian F, Liu H, Huang J. Diffuse optical tomography enhanced by clustered sparsity for functional brain imaging. *IEEE Trans Med Imaging*. 2014;33:2323–31.
3. Lee O, Kim JM, Bresler Y, Ye JC. Compressive diffuse optical tomography: non-iterative exact reconstruction using joint sparsity. *IEEE Trans Med Imaging*. 2011;30:1129–42.
4. Prakash J, Shaw CB, Manjappa R, Kanhirodan R, Yalavarthy PK. Sparse recovery methods hold promise for diffuse optical tomographic image reconstruction. *IEEE J Sel Top Quantum Electron*. 2014;20:6800609.
5. Okawa S, Yoko H, Yamada Y. Improvement of image quality of time-domain diffuse optical tomography with lp sparsity regularization. *Biomed Opt Express*. 2011;2:3334–48.
6. Han D, Tian J, Zhu S, Feng J, Qin C, Zhang B, Yang X. A fast reconstruction algorithm for fluorescence molecular tomography with sparsity regularization. *Opt Express*. 2010;18:8630–46.
7. Candès EJ, Wakin MB. An Introduction to Compressive Sampling. *IEEE Signal Process Mag*. 2008;2:21–30.
8. Beck A, Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J Imaging Sci*. 2009;2:183–202.
9. Figueiredo MAT, Nowak RD, Wright SJ. Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems. *IEEE J Sel Top Signal Process*. 2007;1:586–97.
10. Guo H, Yu J, He X, Hou Y, Dong F, Zhang S. Improved sparse reconstruction for fluorescence molecular tomography with $L_{1/2}$ regularization. *Biomed Opt Express*. 2015;6:1648–64.
11. Zhao L, Yang H, Cong W, Wang G, Intes X. L_p regularization for early gate fluorescence molecular tomography. *Opt Lett*. 2014;39:4156–9.
12. Mohimani H, Zadeh MB, Jutten C. A fast approach for overcomplete sparse decomposition based on smoothed ℓ_0 norm. *IEEE Trans Signal Process*. 2009;57:289–301.
13. Li Y, Santosa F. A computational algorithm for minimizing total variation in image restoration. *IEEE Trans Image Process*. 1996;5:987–95.
14. Hebden JC, Veenstra H, Dehghani H, Hillman EMC, Schweiger M, Arridge SR, et al. Three-dimensional time-resolved optical tomography of a conical breast phantom. *Appl Opt*. 2001;2001(40):3278–87.
15. Tromberg BJ, Shah N, Lanning R, et al. Non-invasive in vivo characterization of breast tumors using photon migration spectroscopy. *Neoplasia*. 2000;2:26–40.
16. Zhu D, Li C. Nonconvex regularizations in fluorescence molecular tomography for sparsity enhancement. *Phys Med Biol*. 2014;59:2901–12.
17. Gao F, Zhao H, Zhang L, Tanikawa Y, Marjono A, Yamada Y. A self-normalized, full time-resolved method for fluorescence diffuse optical tomography. *Opt Express*. 2008;16:13104–21.
18. Vojnovic B. *Advanced time-correlated single photon counting techniques*. Berlin: Springer; 2006.
19. Tehrani JN, Mcewan A, Jin C, et al. L_1 regularization method in electrical impedance tomography by using the L_1 -curve (Pareto frontier curve). *Appl Math Model*. 2011;36:1095–105.
20. Chen W, Wang X, Wang B, Wang Y, Zhang Y, Zhao H, Gao F. Lock-in-photon-counting-based highly-sensitive and large-dynamic imaging system for continuous-wave diffuse optical tomography. *Biomed Opt Express*. 2016;7:499–511.
21. Qin D, Ma Z, Gao F, Zhao H. Determination of optical properties in turbid medium based on time resolved determination. *Proc SPIE*. 2007;6534:65340T.