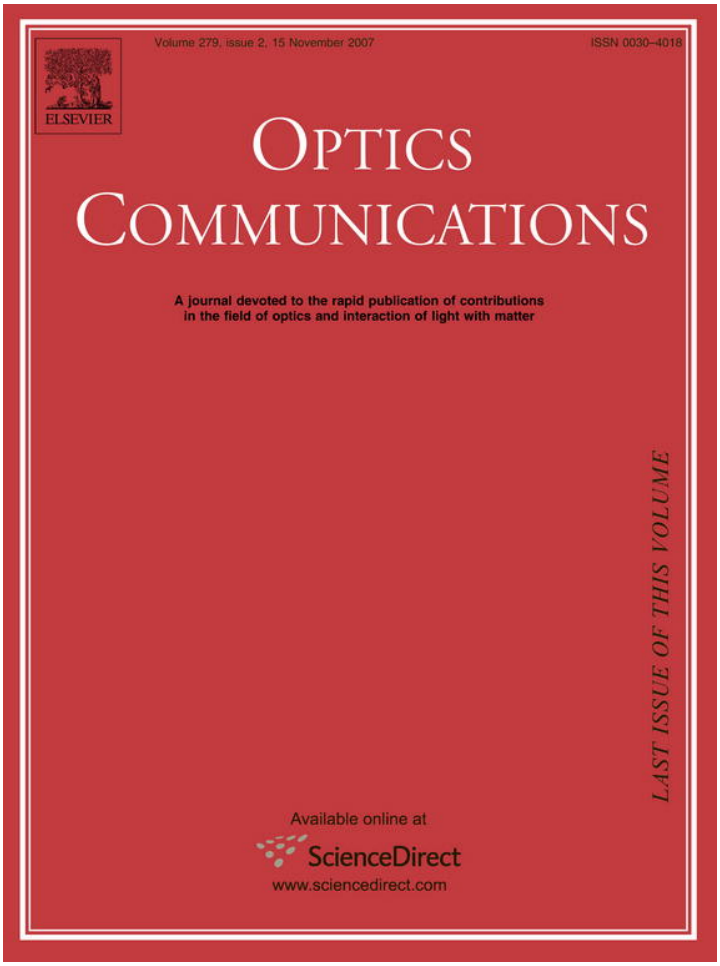

Available online at www.sciencedirect.com
ScienceDirect
Optics Communications 279 (2007) 273–284
OPTICS COMMUNICATIONS
www.elsevier.com/locate/optcom

An efficient method for model refinement in diffuse optical tomography

A.R. Zirak *, M. Khademi [1]

Department of Electrical Engineering, Ferdowsi University of Mashhad, Mashhad, Iran

Received 19 May 2007; received in revised form 23 July 2007; accepted 30 July 2007

Abstract

Diffuse optical tomography (DOT) is a non-linear, ill-posed, boundary value and optimization problem which necessitates regularization. Also, Bayesian methods are suitable owing to measurements data are sparse and correlated. In such problems which are solved with iterative methods, for stabilization and better convergence, the solution space must be small. These constraints subject to extensive and overdetermined system of equations which model retrieving criteria specially total least squares (TLS) must to refine model error. Using TLS is limited to linear systems which is not achievable when applying traditional Bayesian methods. This paper presents an efficient method for model refinement using regularized total least squares (RTLS) for treating on linearized DOT problem, having maximum a posteriori (MAP) estimator and Tikhonov regulator. This is done with combination Bayesian and regularization tools as *preconditioner* matrices, applying them to equations and then using RTLS to the resulting linear equations. The preconditioning matrixes are guided by patient specific information as well as a priori knowledge gained from the training set. Simulation results illustrate that proposed method improves the image reconstruction performance and localize the abnormally well.

Keywords: DOT; Prior information; Bayesian methods; Preconditioner; Regularization; RTLS

1. Introduction

Tomography is imaging by sections or sectioning. The method is used in medicine, archaeology, biology, geology, materials science and other sciences. Traditional tomographic systems such as CT, MRI, SPECT and PET are expensive, use radiation, need special rooms (MRI), and are unable to provide the functional information, and gathering measurements is too slow for the process of tomography [1]. DOT is becoming a useful complement to the current tomographic modalities. It involves the technique of using near infrared (NIR) photons for imaging specific parts of the body to obtain information about tissue abnormalities, such as breast and brain tumors. Although, due to diffusive nature of NIR light, DOT systems have poor spatial resolution, difficult non-linear and ill-posed inverse scattering problems, and contact with the object is necessary, DOT imaging has the advantage of good contrast of functional parameters using a non-invasive and non-ionizing mechanism, good dynamic resolution, up to 100 images/s, and small, even portable, measurement systems [2]. Also, it has a temporal resolution an order of magnitude faster than other new physiological systems such as functional MRI. The reconstructed images of the spatial distribution of tissue parameters can be related directly to physiologically important properties such as blood and tissue oxygenation states. For this reasons, DOT is an interesting and growing research field with a number of medical applications. Its dominant enabling field is related to Bio-Chemical agent detection such as its ability to provide unique information relating metabolic status, and the possibility for non-invasive functional imaging without extrinsic contrast agents by diffusive property of photons in turbid media.

Image reconstruction methods in DOT have been the subject of intense theoretical work and gained considerable momentum over the past decade, marked by the development of two- and three-dimensional imaging techniques,

* Corresponding author. Tel.: +98 21 44063778.
E-mail address: zirak@um.ac.ir (A.R. Zirak).
[1] Associate Professor.

0030-4018/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.optcom.2007.07.054

modeling the photon behavior in media, and different types of numerical schemes to solve it. This modeling is formulated with some approximations [2] which are contaminated by error and can be retrieved with TLS criteria. But, DOT poses a typical ill-posed problem with a large number of unknowns and a relatively limited number of measurements. This necessitates the incorporation of *a priori* information into the inverse problem formulation in order to obtain viable solutions. To tackle the ill-posed nature of the inverse problem and to address the low spatial resolution in DOT, a number of approaches have been developed. Bayesian approach has been suggested to incorporate *a priori* information to the inverse problem formulation [2–8]. Introducing penalty functions and uniform or spatially varying regularization terms within the regularization framework are alternative ways to incorporate *a priori* information into the image reconstruction process [6,9]. However, Bayesian methods are inherently non-linear, such that when using in DOT problems, the TLS criteria is not accessible with traditional methods [10,11]. Most of ways around this difficulty is to retrieve error of model with methods that marginalize model error as noise [2], but they do not retrieve the mismatch between measurements and model, while TLS can do this. Also, error marginalization methods usually results in less image quality.

In agreement with the statistical approach taken in this paper, for better computation and image quality, a new statistically based preconditioners and truncation method are introduced which have MAP estimator property without changing the linearity of the problem. For mismatch removal of model matrix and measurement vector, an RTLS criterion to the resulting preconditioned system is introduced which is not applicable to traditional Bayesian based methods.

This paper is organized as follows: in Section 2, some theories in DOT that are used in this paper are discussed. In Section 3, useful methods for solving linear equations which we are concerned with are recalled. In this section, preconditioning and the regularization theory are reviewed. Section 4 introduces iterative method for RTLS problem and some implementations for selection of preconditioners based on statistical and regularization premises. The details of how we determined the preconditioner and principal components from a library of typical solutions are presented in this section. Simulation results illustrating the performance of RTLS method and preconditioners in DOT problems are presented in Section 5. Section 6 contains the conclusions.

2. Theory

In this section, the simulation basis are described briefly, such as light transport model based on the Rytov approximation of the diffusion approximation and its application to the image reconstruction process. See [1–4] for more detailed derivations.

2.1. Formulation of the problem

DOT is centered around the simple idea that light passes through the body in small amounts, and emerges bearing clues about tissues through which it has passed. One special feature in DOT is the variety of the different measurement or data types that are used for image reconstruction. Three types of systems are widely used in the community: continuous-wave (CW), time-domain and frequency-domain systems [3,4]. The experimental systems in use today utilize either ultrashort input pulses (time-domain systems) or continuous intensity-modulated input (frequency-domain systems). In the latter case, the measurements consist of the complex intensity in terms of the phase shift and amplitude of the transmitted photon density wave, which is a sinusoid modulated continuous wave. The frequency domain version of the problem is the one considered in this paper for simulation, in such a way that the boundary measurements gathered around the boundaries of the object are the measured amplitude and phase of received optical signal with adding 2% zero-mean Gaussian noise to both.

It was recognized that in turbid media such as biological tissue, the radiative transfer equation (RTE) is an accurate model for photon transport using the particle picture of light [3]. Light transport in turbid media can be described accurately by the diffusion approximation (DA) equation, which is an approximation to the RTE.

Let $\Omega \subset R^p$ $(p = 2,3)$ denote the model domain. In the frequency-domain, the DA is expressed by [4,5]

$$-\nabla\cdot(\kappa(r)\nabla\Phi(r,\omega)) + \left(\mu_a(r) + \frac{i\omega}{c}\right)\Phi(r,\omega) = q_0(r,\omega), \quad r \in \Omega, \tag{1}$$

where $\Phi(r,\omega)$ is the photon density at position r and modulation frequency ω (in this work, $\omega = 100$ MHz), $\kappa(r) = 1/3(\mu_a + \mu_s')$ denotes the spatially varying diffusion coefficient with absorption coefficient μ_a and reduced scattering coefficient μ_s', c is the speed of light in the medium, $p = 2,3$ is the dimension of the domain, $q_0(r,\omega)$ is the isotropic source term and $i = \sqrt{-1}$.

To solve Eq. (1) inside the domain under test, the boundary condition must be applied. In the literature, the most frequently used boundary condition is the Robin boundary condition, which is also referred to as the partial current boundary condition [6]. It can be derived as follows: within the DA framework, it can be shown that the total inward directed photon at a point r on the boundary is zero. In this paper, a Robin (*Type III*) boundary condition is used, which is given as follows [6,7]:

$$\Phi(r,\omega) + \frac{1}{2\gamma}\kappa(r)\frac{\partial\Phi(r,\omega)}{\partial v} = g(r,\omega) \quad r \in \partial\Omega,$$

where $g(r,\omega)$ models the boundary sources, γ is a dimension-dependent constant ($\gamma_2 = 1/\pi$, $\gamma_3 = 1/4$) and v is the outer normal at boundary domain, $\partial\Omega$.

2.2. Image reconstruction in optical tomography

The image reconstruction problem is to estimate the optical properties within the object when the amount of guided light and the actual measurement on the boundary of the object are given. Image reconstruction in DOT always done by sequential solving of two problems, forward and inverse problems [8,9]. Roughly speaking, the forward problem in DOT evaluates the photon propagation in tissue and is usually formulated by the DA equation. The inverse problem consists in retrieving the spatially varying image of the object by comparing predicted and actual measurements. In this paper, the forward problem is solved with finite element based methods (FEM). The key principle in the FEM is the reduction of the general, continuous, problem to one of matrix algebra of finite size. Each node in the FEM mesh is labeled according to the region obtained from simulated MRI images [12].

The inverse problem solving methods are divided into two classes, which are the linear (difference or perturbation) methods and the methods based on the non-linear optimization approach. In the former category, measurements are taken before and after a small perturbation in the optical properties. Consequently, if $\Im(\mu_a, \mu_s)$ is the FEM based forward operator for light transport which maps the spatial absorption and scattering parameters to the predicted measurements, the relation of perturbation in optical properties to perturbation in optical measurement can be linearized with the perturbation equation by [13]:

$$\Delta y = J\Delta x, \tag{2}$$

where $\Delta y = y - \Im(\mu_a, \mu_s)$ is the difference between actual and predicted measurement, respectively, and $\Delta x = (\mu_a, \mu_s) - (\mu_{a,ref}, \mu_{s,ref})$ is the difference between estimated and approximated images. The measurements in (2) are complex, and inversion would lead to a complex parameter update. For this reason we split the measurement vector into real and imaginary parts with a commensurate splitting of the linearized derivative operators. In addition, when considering log of the data the splitting associates the real part with logarithmic amplitude, and the imaginary part with phase [9]. The matrix J is named the Jacobian or sensitivity matrix for the forward model, which is the first Frechet derivative of $\Im(\mu_a, \mu_s)$ with dimension $MS \times 2N$ where M is the number of measurements, S is the number of sources (MS rows or equations) and N is the number of FEM based nodes into which the object is discretized ($2N$ column or scattering and absorption parameters) [13]. To yield Eq. (2), We have employed the perturbation approach with a first-order Rytov approximation [7] to solve the forward problem in the frequency domain.

For non-linear optimization methods, the regularized least squares problem is to estimate absorption and scattering distributions which minimize the objective function

$$\Psi = \|(y - \Im(\mu_a, \mu_s))\|^2 + \Re(\mu_a, \mu_s) \tag{3}$$

where $\|\cdot\|$ denotes the Euclidean norm and $\Re$ is the penalty operator. Further, the first term in objective function (Ψ) is least squares based for data fitting and the positive second term to be minimized, penalizes unexpected growth of the solution in respect to background, which is named a regularizing penalty. The non-linear minimization problem (3) is usually solved iteratively using gradient methods such as the non-linear conjugate gradient [14,15] or Newton based methods [9] to minimize the objective function over the search space of optical parameters. Newton based methods converges faster owing to the contribution of higher order information. However, the price paid is the reduction in robustness, i.e. it is more sensitive to poor initial estimates. To overcome this, a Gauss–Newton approach [9] to the inverse solver in optical tomography is used. In the Gauss–Newton method, after computing Jacobian matrix ($J_{(i)}$) and predicted measurements ($\Im(\mu_a^{(i)}, \mu_s^{(i)})$) with the latest estimate of optical properties that are computed at iteration i, the image can be updated by solving the minimization problem

$$\begin{pmatrix} \mu_a \\ \mu_s' \end{pmatrix}_{(i+1)} - \begin{pmatrix} \mu_a \\ \mu_s' \end{pmatrix}_{(i)} = c_{(i)} \left(J_{(i)}^T W J_{(i)} + \frac{1}{2} H_B^{(i)} \right)^{-1} \left(J_{(i)}^T W \left(y - \Im(\mu_a^{(i)}, \mu_s'^{(i)}) \right) - \frac{1}{2} g_B^{(i)} \right), \quad i = 1, 2, \ldots \tag{4}$$

iteratively, where i is the iteration index, μ_a, $\mu_s \in R^{N/2}$, W is a positive definite covariance matrix, $c_{(i)}$ is a step size parameter, and $g_B^{(i)}$ and $H_B^{(i)}$ are gradient (first derivative) and Hessian (second derivative) of the penalty $\Re(\mu_a, \mu_s)$, respectively.

It is shown that [8], the linear methods has a good performance such as robustness for poor initial estimates in the first stage of the iterative process. But in the final stages of the iterations, where the estimate of optical property is closed to the actual ones, Gauss–Newton method is very good because of its quadratic convergence and low oscillation during convergence [9]. For this reasons, we have used the linear deference method (2) at the first and the Gauss–Newton method (4) at the final stage.

3. Traditional methods for solving linear equations

In order to focus on numerical aspects of Eqs. (2) and (4), they can be represented in the matrix form by concerning with the solution of large linear equation systems contaminated by additive Gaussian noise:

$$Ax = b + e, \quad A \in R^{m \times n}, \quad x \in R^n, \quad b, e \in R^m, \tag{5}$$

where e is the additive noise vector, and b and A are named the measurement vector and model matrix, respectively. In DOT the actual measurement of each photon is most often corrupted by shot noise statistics, which originates from Poisson statistics. However, With a sufficiently large number of detected photons, as the intensity tends to infinity,

the probability distribution of the shot noise is normally considered in the Gaussian limit [7]. So, e in (5) is assumed to have Gaussian distribution, where its variance is expected to be proportional to the number of photons at the detector and spatially uncorrelated for simplicity.

In particular, we are concerned about the solution of linear discrete ill-posed problems for which the corresponding model matrices A are of ill-determined rank, contaminated by noise, and as large as to make its factorization impossible and not explicitly available. Besides, for a relatively small solution space, the system is considered overdetermined and so model matrix space and measurement vector are not coincident. In this section we extend the discussion into three major and popular methods for the conversion of linear systems to have a better behavior and solutions. Through this methods, one that is used frequently in solving of DOT problems is Tikhonov regularization.

3.1. Preconditioning theory and dimension reduction

In traditional iterative reconstruction methods, preconditioning is a technique which improves the convergence rate by transforming the matrix A into a new matrix with more desirable spectral characteristics [16]. With this perspective, the equation

$$(L_1AL_2^{-1})(L_2x) = L_1b \tag{6}$$

is named the preconditioned system of Eq. (5), which has the same solution. L_1 and L_2^{-1} are symmetric positive definite matrices which are named left and right preconditioners, respectively, such that the condition number of the resulting matrix $L_1AL_2^{-1}$, in general, will be smaller. Naturally, the closer resulting matrix is to the identity, the faster we can expect an iterative method to converge. So, the dimension subspace of the problem can be reduced effectively with preconditioning [16].

This set of transformations can be directly applied to formulate preconditioned iterative algorithms such as the preconditioned conjugate gradient. After the last iteration, the image can be recovered by applying the transformation. Furthermore, the side where the preconditioner is applied is very important. In the iterative solution of linear discrete ill-posed problems a right preconditioner is closely related to available or inferred information about the solution, while a left preconditioner conveys information about the noise in the data or model matrix whose statistical properties may be known. A survey of popular preconditioning strategies for linear systems iterative solvers can be found (e.g. in [16]).

3.2. Regularization of ill-posed problems

When the linear Eq. (5) comes from the discretization of a diffusion approximation, many of the singular values of the coefficient matrix A are very close to the origin, so the ill-conditioning of the coefficient matrix for these linear systems is typically very large [17]. Quite often, one of the following results can occur in the evaluation of the goal of the inverse problem: (1) the solution does not exist, (2) the solution is not unique, or (3) solving for the solution is not stable, i.e. a tiny perturbation (error) in b (measurements) will be amplified so that it results in a large perturbation in x (image). If any of the above results occurs, the inverse problem is said to be ill-posed (in the Hadamard sense) [18] sometimes so much as to make the computed image useless. An ill-posed problem must be converted into a well-posed version in order to be solved. The technique for this conversion is called regularization. For a linear ill-posed problem, the following methods are among the most commonly used which all of them are satisfied with proposed method: (1) the truncated methods such as truncated SVD (TSVD), (2) Tikhonov regularization and (3) the truncated total least squares (TTLS).

To solve a linear ill-posed problem using the TSVD, the SVD of A must be computed and then the small singular values which correspond to cheap subspaces must be omitted. the suitable subspace called coarse space. This is not realistic when there is not sufficient gap between large and small singular values of A, or when A is very large since the most efficient algorithms for computing the SVD of a general matrix has computational complexity $O(N^3)$. For more details see [19,20].

An alternative to TSVD is Tikhonov regularization. Tikhonov regularization replaces the linear system (5) with the minimization problem

$$x_{\text{tikhonov}} = \arg\min_x \{\|Ax - b\|^2 + \lambda\|L(x - x_0)\|^2\} \tag{7}$$

which the second term in the functional to be minimized penalizes unexpected growth of the solution. The Matrix L decides how the computed solution should be allowed to grow and the regularization parameter λ balances the effect of the prior with predicted-actual measurements mismatch. Note that for a general case, i.e. without the *a priori* information, L becomes identity matrix. To compare proposed technique with traditional Tikhonov method, in the later, L and λ are calculated as [9,18]:

$$(L^{\mathrm{T}}L)_{ij} = \begin{cases} nn & \text{if } i = j \\ -1 & \text{if } j \text{ is neighbour of } i \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

where nn is the number of the neighbors of basis component i in FEM. We have used the so-called *L-curve*, which has been widely used to choose an optimal regularization parameter λ for solving Eq. (7). Details of implementing the L-curve to choose λ can be found in [21]. Tikhonov regularization is one of the most popular methods in DOT. Moreover, the computational complexity for solving optimization problem (7) is only on the order of $O(N^2)$ [8,14].

A third method for regularizing linear ill-posed problems is the TTLS [10]. It can guarantee the existence of the solution by adapting the measurement space with the model space. The TLS method can produce a robust solution only in linear systems such as Eq. (5) when the A

matrix and b are both contaminated with noise whereas other methods only consider the noise in b. For more detail see [11].

3.3. Regularized total least square method

Total least squares (TLS) is a method for treating an overdetermined system of linear equations $Ax \approx b$, where both the matrix A and the vector b are contaminated by noise. The inequality may often be attributed to two possibly distinct mechanisms: model mismatch and measurement inaccuracies. Model mismatch encompasses inaccuracies for discretization even when exact measurements are used. Measurement inaccuracies, on the other hand, account for possible deviations of the measured data from the true (unknown) data (in the absence of model mismatch). The TLS estimation of x is obtained by [10]:

$$\min \|(E,f)\|_{\mathrm{F}} \quad \text{subject to } (A+E)x = b+f,$$

where both E and f must to be minimized with the *Frobenius* norm $\|\|_{\mathrm{F}}$. The best method for solving the above equation, is based on optimization, which leads to:

$$x_{\mathrm{TLS}} = \mathrm{argmin}_x \frac{\|Ax-b\|^2}{1+\|x\|^2},$$

Also, in practical situations, the linear system is often ill-conditioned. These happen when the system is obtained via discretization of ill-posed problems such as DOT when using Newton based methods. In these cases the TLS solution can be physically meaningless and thus regularization is essential for stabilizing the solution. For this reason, RTLS was addressed by several approaches such as truncation methods and Tikhonov regularization [16,10,11]. In this paper RTLS is applied to all of the preconditioned linear equations as the minimization problem of [10]:

$$x_{\mathrm{RTLS}} = \arg\min_{x\in R^n} \left\{ \frac{\|Ax-b\|^2}{\|x\|^2+1} : \|L_{\mathrm{s}}x\|^2 \leqslant \rho \right\}, \tag{9}$$

where $\rho > 0$ is a regularization parameter and $L_{\mathrm{s}} \in R^{k\times n} (k \leqslant n)$ is a regularization matrix that defines a (semi)norm on the solution which is frequently chosen to approximate the first or second derivative operator.

4. Implementation

In this paper, preconditioners are introduced for simplification of regularization and MAP estimation without changing the linearity of systems of type (5). So, applying some criterions such as reduction model with error marginalization and accelerating convergence of iterative algorithm can be simple and realizable. Besides, by using TSVD to statistical information matrices, the solution space is reduced to minimum complexity. For this reason, the approach is efficient for simplification and application of three kinds of regularization methods which are discussed in (Section 3.2). From the point of view of regularization, the traditional preconditioners which only improve the speed of convergence seem to be of little use in DOT. Indeed, accelerated convergence by preconditioning may lead to an iterative method where the noise takes over immediately and the regularization property is lost. It has been demonstrated that if the preconditioner is selected properly, it improves the quality of the iterative solution [22]. Indeed, the rule of thumb is that a matrix L that works well in Tikhonov regularization (7), works also as a preconditioner. In this work, this property is investigated from the point of view of Bayesian statistics. For this reason, a good preconditioner in DOT is constructed from the Cholesky factor of the inverse of the prior covariance matrix of the unknown. Such preconditioners are referred as *priorconditioners*.

Since, in medical applications, a prior density for the unknown might be replaced by a large collection of realizations, useful priorconditioners can be constructed by sampling. For this reason, we suppose that a learning set of typical vectors x (anatomical information or data collected by previous surgical interventions) and the corresponding measurement b are available. The *priorconditioners* are then constructed from the sample-based estimates of the expected value and covariance matrix of the unknown. By remarking anomaly with the statistical interpretation, we have stabilized the computation and regularize the covariance matrix, which is further discussed later in the rest of the paper.

4.1. Statistics and regularization as preconditioner perspective

DOT has a variety of explicit possible clinical applications specially in physiological studies, for example the imaging functional activity of the brain region [2,23]. But, it is well-known that traditional DOT reconstruction algorithms do not produce satisfactory reconstructions when applied to sparse projection data or the model matrix [24]. In recent years, many papers present and review results suggesting that statistical inversion methods can be successfully used for reconstruction in DOT.

Although many statistical methods have been proposed for the restoration of tomographic images, their use in medical environments especially in functional imaging, has been limited due to two important factors. These factors are the need for greater computational time than deterministic methods and the selection of the hyperparameters in the image models. For this manner, we consider the compatibility of prior assumptions of regularization scheme and the actual prior information on optical property distribution from a Bayesian perspective and then simplification of applying them as preconditioners to linear systems, and estimation of hyperparameters with aide of typical Sequences.

In this paper, random variables are denoted by capital letters and their realizations are denoted by lowercase letters. In the Bayesian approach the optical property distri-

bution X and the measured voltages Y are assumed to be multivariate random variables with some joint probability density $p(x,y)$. In inverse problems, all variables are interpreted as random ones, the randomness really reflecting our lack of information about their values. Hence, instead of the deterministic Eq. (5), we consider its stochastic extension,

$$AX = B + E, \quad A \in R^{m\times n}, \quad X \in R^n, \quad B, E \in R^m, \tag{10}$$

where X, B and E are random variables instead of x, b and e in deterministic model, respectively, and A is the deterministic model matrix. Most papers use the additive noise model E, but the ideas of this paper are more generally applicable, so that the model error in A can be added to E with good approximation. $p(x)$ denotes the *prior* probability density of image X, which expresses the degree of information about the values of X prior to measuring B. The *likelihood* density, denoted by $p(b|x)$, is the probability density of B given the realization $X = x$. The probability density of X given $B = b$ is called the *posterior* density and is denoted by $p(x|b)$;

$$p(x|b) = \frac{p(x)p(b|x)}{p(b)} \tag{11}$$

This is the solution of the inverse problem (10) in the Bayesian frame of mind. Based on the posterior density, we may define various estimates of the image x. The most commonly used statistical estimates are the *conditional mean* (CM) and MAP estimates,

$$x_{\mathrm{MAP}} = \arg\max_x p(x|y), \quad x_{\mathrm{CM}} = \int x p(x|y)\,\mathrm{d}x,$$

provided that such estimates exist. Consider the linear additive noise model (10) under the assumption that X and E are mutually independent Gaussian random variables with $X \sim N(\mu_x,\Gamma_x)$ and $E \sim N(0,\Gamma_e)$, that is, the random variable X has mean $\mu_x \in R^n$ and its covariance matrix $\Gamma_x \in R^{n\times n}$ is symmetric, positive definite, and E is zero mean Gaussian noise with covariance $\Gamma_e \in R^m$. Then, Bayes' formula implies that the posterior density is

$$\begin{aligned} p(x|b) &\propto \exp\left(-\frac{1}{2}[(x-\mu_x)^{\mathrm{T}}\Gamma_x^{-1}(x-\mu_x) + (b-Ax)^{\mathrm{T}}\Gamma_e^{-1}(b-Ax)]\right) \\ &= \exp\left(-\frac{1}{2}[\|L_x(x-\mu_x)\|^2 + \|L_e(b-Ax)\|^2]\right) \\ &= \exp(-\Psi(x|b)) \end{aligned} \tag{12}$$

where the Cholesky factors of Γ_x^{-1} and Γ_e^{-1} are denoted by L_x and L_e, respectively, i.e., $L_x^{\mathrm{T}}L_x = \Gamma_x^{-1}$, with L_x upper triangular. Under these assumptions, the maximum *a posteriori* estimate x_{MAP} coincides with the conditional mean estimate x_{CM}, which is the centre point of the posterior density given above, and they are the solutions to the minimization problem

$$x_{\mathrm{MAP}} = x_{\mathrm{CM}} = \arg\min(\Psi(x|b)).$$

For solving this minimization problem, first the new random variable is introduced:

$$W = L_x(X - \mu_x),$$

which is a Gaussian white noise $W \sim N(0,I)$, since

$$\mu_W = E\{W\} = 0, \quad \Gamma_W = E\{WW^{\mathrm{T}}\} = L_x\Gamma_x L_x^{\mathrm{T}} = I,$$

where I is the identity matrix. Therefore, the upper triangular matrix L_x is a *whitening matrix* for X and owing to the above properties of W, the resulting equation has a better convergence property. For a given realization of the image x, if we define $b_0 = b - A\mu_x$, we can write

$$\Psi(x|b) = \|w\|^2 + \|L_e(b_0 - AL_x^{-1}w)\|^2 = \left\| \begin{bmatrix} L_eAL_x^{-1} \\ I \end{bmatrix} w - \begin{bmatrix} L_eb_0 \\ 0 \end{bmatrix} \right\|^2 \tag{13}$$

Then, the MAP estimator is the solution of this linear system in the least square sense, in the context of iterative solvers with appropriate regularization:

$$L_eAL_x^{-1}w = L_eb_0, \quad w = L_x(x - \mu_x) \tag{14}$$

Since measurements in ill-posed problems in DOT are always noisy, we can also use information about the statistics of the noise in the solver to improve the quality of the computed solution. This is well known where whitening of the noise is often a preprocessing step for an inversion algorithm. In the iterative linear solvers framework, the whitening of the noise naturally defines a left preconditioner.

The Gaussian MAP estimate (and so the linear Gaussian statistical model) is intimately related to the classical Tikhonov regularization of ill-posed problems. Indeed, if it is assumed that the additive noise is white Gaussian, i.e., $\Gamma_e = \sigma^2 I$, the MAP estimate agrees with the Tikhonov regularized solution:

$$\lambda = \sigma^2, \quad L = L_x, \quad \mu_x = x_0.$$

4.2. Iterative methods

Beginning with an initial optical property estimate, the forward model needs to be evaluated to produce the predicted measurement. An update of the optical property is then computed from the difference between the predicted and actual boundary measurements utilizing the gradient information for updating the previous optical property and the evaluation of the forward model for a new prediction of measurement. This process is repeated until the satisfactory match is found between the predicted and actual measurements. In order to retrieve the spatial distribution of optical properties (or image) of the unknown media, one needs to solve an ill-posed non-linear optimization problem which is best addressed using iterative gradient-based approaches. The solution of linear discrete ill-posed problems by iterative methods equipped with a suitable stopping rule has gained a lot of attention in recent years, in view of its speed and ease of implementation [25].

Among the various iterative methods proposed in the literature, the conjugate gradient (CG) method for the symmetric positive definite case, the generalized minimal residual (GMRES) method for the square non-symmetric case, and the conjugate gradient for least squares (CGLS) method and its implementational variations for the general non-square case have been the methods of choice for linear discrete ill-posed problems.

For the iterative solution of preconditioned linear Eq. (14), it only requires to replace b, A and x in Eq. (5) with $L_e b_0$, $L_e A L_x^{-1}$ and w, respectively, but when we use preconditions, the equations have a better solution when applying GMRES with respect to CG based methods.

The use of preconditioners to accelerate the rate of convergence of these iterative methods is quite prevalent. But the choice of preconditioners would be more efficient to avoid accelerating the convergence of those components dominated by amplified errors, when applied to the linear discrete ill-posed problems. For this reason, from a decade ago, preconditioners for iterative methods which aim at accelerating only the convergence of the part of the spectrum associated with the signal, while leaving the portion associated with the noise alone, were first proposed in linear algebra [14]. All preconditioners of this type, which are related to truncated SVD, need to somehow decompose the spectrum of matrix A to separate the eigenvalues associated with the signal from those associated with the noise. This task is easy when the spectrum of A shows a well-marked gap between larger and smaller eigenvalues, while it becomes very difficult in DOT. Because most of small eigenvalues belong to important information of image such as small tumors, as well as the eigenvalues or singular values decrease smoothly to zero. For comparison to our proposed method, we have used this kind of preconditioners.

In a different approach, when we use priorconditioners in (14), preconditioning is viewed as a tool to improve the quality of the computed solution rather than to accelerate the convergence. In view of this observation, which is motivated by the reformulation of $\Psi(x|b)$ in terms of w, lead to the system (14), we use the factor L of Γ^{-1} as a right and left preconditioners. Our simulations indicate that it is not often necessary to take L to be the Cholesky factor of the inverse of Γ. In fact, if we have a factorization of Γ of the form $M^{\mathrm{T}}M = \Gamma$, we may choose $L = M^{-1}$. Note that since the construction of the priorconditioner uses no information about matrix A, we do not expect it to affect its spectral distribution. On the other hand, since the priorconditioner carries a lot of information about the distribution of solution X, once it has been computed it may be used in connection with several different linear systems. Finally, it is easy to see that priorconditioners can be used in combination with standard preconditioners for faster convergence.

In this section, we have assumed that the prior density of X is directly given as a Gaussian density. however, the prior density is often non-Gaussian and, in fact, the *a priori* knowledge about the solution may not be formulated at all in statistical terms, but rather as a large database of sample solutions. For *termination criteria* definition, a popular method that is used in iterative solution of linear systems, and the one that we used here, are referred to as a truncated iteration. The idea behind regularization by a truncated iteration is that in the first few iterations the computed solution approaches the exact solution, but as the iterations continue, amplified noise components start to dominate the computed solution. Therefore, to make iterative methods suitable for the solution of linear discrete ill-posed problems, it is necessary to equip them with suitable termination criteria which stop the iteration prior to the inclusion in the computed solution of amplified noise components.

4.3. Sample-based priors

A central challenge in statistical modeling of inverse problems is constructing informative and reliable prior densities. In this paper, we discuss sample-based estimation of the prior. Sample-based priors have been discussed in [26]. Another equally important question concerning priors is how to avoid typical priors that are biasing towards a reasonable, but incorrect, solution. In particular, in medical imaging, the prior should favor the typical or normal solutions that we expect to see, but at the same time it should allow the appearance of abnormalities or anomalies that are often of central interest. This issue has been previously addressed in [27–29]. These two related questions will be addressed in this section. Our approach to all these issues is based on random sampling.

4.3.1. Computing of priors with random sampling

Assume that we have access to a sample of realizations of the random variable X, as well as of the corresponding set of data. this sample of pairs are named a *training set* and denoted by

$$\vartheta_0 = \{(x_1, b_1), (x_2, b_2), \ldots, (x_N, b_N)\}, \quad x_j \in R^n, \ \ b_j \in R^m$$

with $N \geqslant n$. The noise level of the measurements may or may not be considered in sampling. If the prior is Gaussian or another standard parametric distribution, efficient random vector generators can be used. More generally, the sample can be generated by using Markov chain Monte Carlo (MCMC) techniques. In DOT, the training set could consist, e.g., of previous measurements b_j combined with information obtained by surgical or experimental interventions, or it could have been generated, e.g., by using a computational anatomical or physiological model such as previous MRI or CT image. This set is used to set up a prior model that is adjusted to the estimation method of choice. Here, we assume that the vectors x_j are represented as discretized approximations. We start by setting up a prior model based on the training set. If N is large and the sample is representative, it is possible to estimate the probability density of the underlying variable X using this sample. We must note that if a sample of true realizations

is given, they may not distributed normally, But for constructing priorconditioners, we must seek a Gaussian approximation of the prior density. The Gaussian distributions are completely characterized by the second-order statistics. Based on the available sample, the sample mean ($\tilde{\mu}_x$) and the sample covariance ($\tilde{\Gamma}_x$) which are estimation of real mean and covariance obtained as

$$\begin{aligned}\tilde{\mu}_x &= \frac{1}{N}\sum_{j=1}^{N} x_j \approx E\{X\} = \mu_x,\\ \tilde{\Gamma}_x &= \frac{1}{N}\sum_{j=1}^{N} x_j x_j^{\mathrm{T}} - \tilde{\mu}_x\tilde{\mu}_x^{\mathrm{T}} \approx E\{XX^{\mathrm{T}}\} - \mu_x\mu_x^{\mathrm{T}} = \Gamma_x\end{aligned} \tag{15}$$

The higher order moments can be estimated as well, and they can be used to assess the fidelity of the Gaussian approximation.

Always, the prior favor the typical solutions so that resists the appearance of abnormalities or anomalies that are often of central interest. The more severe problem, in particular, from the point of view of DOT applications, is sample's inability to reproduce outliers. Assume, for instance, that the training set represents head intersection images. The major part of the images corresponds to normal skull and brain, while the few outliers representing anomalies, such as tumors, that might be in the set have a negligible effect in the averaging process. As a consequence, the anomalous features will not be represented by the PCA subspace vectors, when the purpose of the imaging process might have been, in fact, to detect these anomalies.

To overcome the aforementioned shortcomings, we propose a stochastic model that accounts for both the *regular* and *anomalous* parts. We write a model

$$X = X_{\mathrm{r}} + X_{\mathrm{a}}, \tag{16}$$

where X_{r} and X_{a} are stochastically independent random variables. The *regular part* X_{r} is assumed to have mean and covariance μ_{r} and Γ_{r}, respectively, calculated from the training set. The *anomalous part* X_{a} accounts for all features not captured by regular part or the PCA reduced model that is proposed in the next. To define the statistics of the anomalous part, we may use any prior information that we have concerning the anomalies that we expect to encounter. For instance, if we expect smooth anomalies of a given size, we use a smoothness prior with a properly chosen correlation length. In this work, we shall assume that X_{a} has zero mean and that its covariance (Γ_{a}) is defined up to a multiplicative constant. Hence, the second-order statistics of X is of the form

$$E\{X\} = \mu_x = \mu_{\mathrm{r}}, \quad \mathrm{cov}(X) = \Gamma_x = \Gamma_{\mathrm{r}} + \alpha\Gamma_{\mathrm{a}}, \tag{17}$$

for some $\alpha > 0$. Assuming that the resulting covariance matrix Γ_x is invertible, we may now write a whitened version of the estimation problem, as (14). The parameter α plays an important role in three different ways. First, even when the correlation structure of the anomaly may be known, the relative anomaly variance, compared to the regular part, may be unknown. Secondly, it is important to guaranty that the stochastic model (16) does not produce false artifacts. Therefore, we adjust the parameter α using the training set so that the prior model performs well on the training set ϑ_0. Lastly, from the numerical point of view, α can be seen as a regularization parameter that guarantees that Γ_x is numerically positive definite, hence ensuring the existence of its Cholesky decomposition. We remark that often sample-based covariance matrices are rank-deficient, because the samples collection is either not sufficiently large or the samples are not independent. To avoid the problem of working with a singular matrix, we add to the sample-based covariance a small multiple of the identity. The effect of this form of regularization of the covariance is not only to make the matrix invertible, but also to allows for anomalies in the solution to appear, which could not be captured from the available sample.

4.3.2. PCA reduced model

In medical applications, the vectors x_j represent typical features of the random variable X. For this reason, the vectors can not be very dissimilar. Consequently, the space spanned by the realizations may be a proper subspace even if $N \geqslant n$, and hence Γ_x is rank deficient or of ill-determined rank with satisfactory approximation of its real value. Without loss of generality, also in linearized perturbation DOT problems, we may assume that the mean of X vanishes. Introduce the singular value decomposition of the matrix Γ_x,

$$\Gamma_x = VDV^{\mathrm{T}}, \quad V = [v_1, v_2, \ldots, v_n], \quad D = \mathrm{diag}[d_1, d_2, \ldots, d_n],$$

where the orthonormal singular vectors v_j correspond to the singular values d_j such that we have; $d_1 \geqslant d_2 \geqslant \cdots \geqslant d_r > d_{r+1} = \cdots = d_n = 0$. In practice, we identify with zero those singular values that are smaller than a given threshold value, that can be thought of as the working precision. We write

$$V_0 = [v_1, \ldots, v_r], \quad V_1 = [v_{r+1}, \ldots, v_n],$$

and further, we split X in parts as

$$\begin{aligned}X &= V_0(V_0^{\mathrm{T}}X) + V_1(V_1^{\mathrm{T}}X) = V_0X_0 + V_1X_1,\\ &X_0 \in R^r, \ X_1 \in R^{n-r}.\end{aligned} \tag{18}$$

We have

$$E\{\|X_1\|^2\} = \mathrm{trace}(E\{X_1X_1^{\mathrm{T}}\}) = \mathrm{trace}(V_1^{\mathrm{T}}\Gamma_xV_1) = \sum_{j=r+1}^{n} d_j = 0,$$

in other words, $X = V_0X_0$ with probability 1. Therefore, if we trust the covariance matrix Γ_x, we may write a *reduced model*

$$B = AX + E = AV_0X_0 + E, \quad AV_0 \in R^{m\times r}, \tag{19}$$

and the model reduction error has zero probability of occurrence. The above model is equivalent to the *principal component analysis (PCA) model* [22] which has a capabil-

ity that it is sufficient to compute the SVD of covariance matrix only once for analogous problems whereas if we use A matrix for model reduction, we must calculate TSVD of it for any iteration. Simulation results shows that the proposed PCA model in DOT has a better image quality with respect to traditional TSVD. In addition, if we define

$$W_0 = D_0^{-1/2} X_0 = D_0^{-1/2} V_0^{\mathrm{T}} X \in R^r,$$

where $D_0^{-1/2} = (D_0^{1/2})^{-1}$ and $D_0^{1/2} = \mathrm{diag}[d_1^{1/2}, d_2^{1/2}, \ldots, d_r^{1/2}]$ we observe that:

$$E\{W_0 W_0^{\mathrm{T}}\} = D_0^{-1/2} V_0^{\mathrm{T}} \Gamma_x V_0 D_0^{-1/2} = I_{r\times r},$$

i.e., W_0 is r-variate white noise. The *whitened PCA model* for solving X can be written as

$$L_e B = L_e A V_0 D_0^{1/2} W_0 + L_e E, \quad X = V_0 D_0^{1/2} W_0. \tag{20}$$

The matrix $V_0 D_0^{1/2} \in R^{n\times r}$ acts as a whitening preconditioner. Observe that the solution is automatically in the subspace spanned by the eigenvectors $v_1, \ldots, v_r$. Hence, the PCA model is a *reducing constraint*, since it forces the solution to a low-dimensional space. The clear advantage of the PCA (19), or its whitened version (20), is that it takes full advantage of the prior information, and when $r \ll n$, the degree of ill-posedness may decrease, and so does the required computational work. also, when applying only informative components in covariance matrix, the important parts such as anomalous dose not destroyed in reconstruction process. however, in medical applications, there are clear disadvantages. The first problem is related to some approximations such as (15) and ignoring some eigenvalues and their corresponding eigenvectors. In general, we do not know to what extent the training set is a sufficient sample. However, the error with given N may have a significant effect on the eigenvalues that are used to determine the truncation parameter r. also there may be some eigenvalues that converge smoothly to zero.

4.4. Applying RTLS to preconditioned PCA model

The RTLS problem was extensively studied in recent years, but A key difficulty with this problem is its non-convexity in DOT problems that not guarantee to converge to a global optimum. To avoid this difficulty, when solving linearized DOT problems, we used globally and efficiently convergent algorithm which is proposed in [11]. Here we used a practical iterative algorithm utilizing the shifted inverse power method for which the only prescribed parameter, ρ in (9), should be determined *a priori* from knowledge of the underlying physical model. The algorithm then determines both the Lagrange multipliers and estimates a RTLS solution concurrently. The only assump-

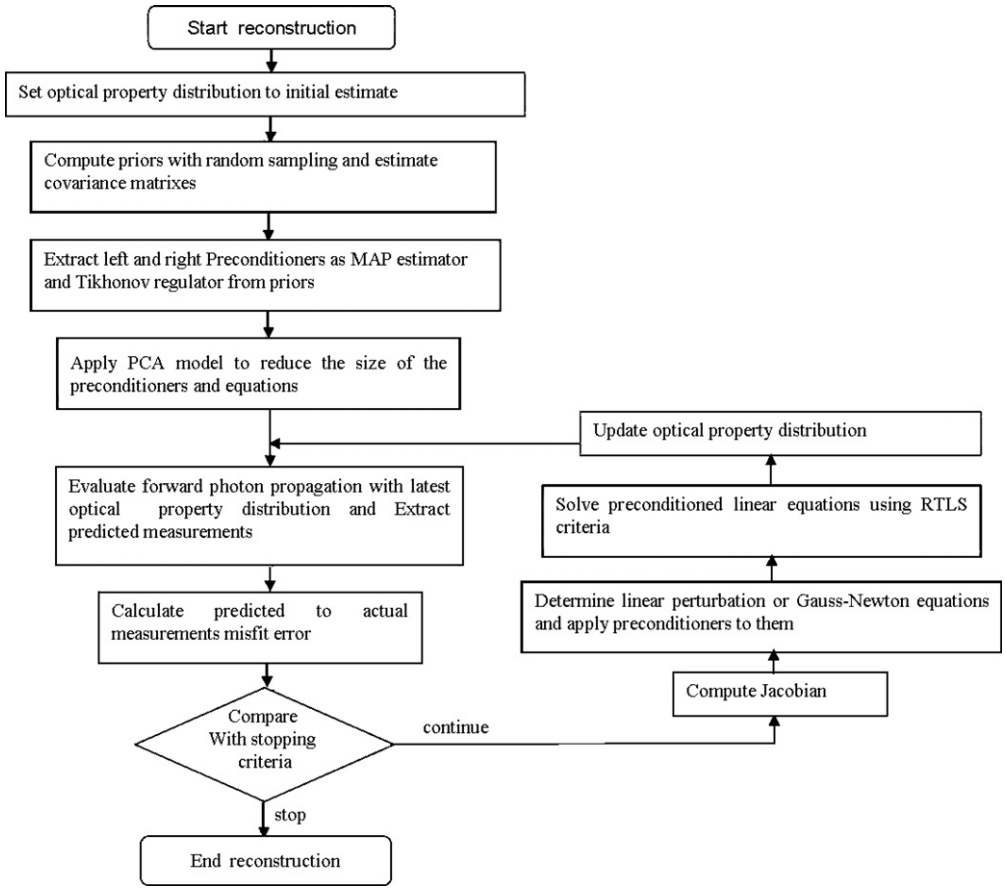

Fig. 1. Flow chart of the proposed reconstruction method.

tion made is that the problem must be well defined. This is achieved when using proposed preconditioners and PCA methods. In this paper, RTLS is applied to whitened PCA model (17). Also, we used Cholesky factor of data covariance matrix, L_s in (9), as a regularization matrix to RTLS problem.

4.5. Flow chart of the proposed algorithm

Based on the proposed methods in Section 4, a flow chart is drawn in Fig. 1 to illustrate the detailed computational steps for a complete reconstruction. From the flow chart, five proposed key steps can be identified: (a) computation of priors with random sampling, (b) extraction of preconditioners from priors, (c) applying reduced PCA model to the right preconditioner, (d) updating the reconstruction equations by applying preconditioners to them and (e) solving preconditioned linear equations using RTLS criteria. The accuracy of these procedures are essential for a successful reconstruction.

5. Numerical results

In this section, we provide numerical examples to validate the applicability of the proposed preconditioners.

5.1. Simulated test phantom

Two multi-layered phantoms are simulated with different absorption and scattering coefficients in a circular object of diameter 100 mm and infinite height. A two-dimensional cross section of these phantoms are shown in Figs. 2a and 3a in Fig. 2a, three regions are shown, region 0 ($\mu_a = 0.1\ \text{mm}^{-1}$ and $\mu_s' = 10\ \text{mm}^{-1}$) and region 1 ($\mu_a = 0.02\ \text{mm}^{-1}$ and $\mu_s' = 2\ \text{mm}^{-1}$) with typical prior information and finally region 2 ($\mu_a = 0.08$ and $\mu_s' = 8\ \text{mm}^{-1}$) which is simulated as a anomaly with few prior information. These values in Fig. 3a are ($\mu_a = 0.015\ \text{mm}^{-1}$ and $\mu_s' = 1.5\ \text{mm}^{-1}$), ($\mu_a = 0.15\ \text{mm}^{-1}$ and $\mu_s' = 15\ \text{mm}^{-1}$) and ($\mu_a = 0.01$ and $\mu_s' = 1$), respectively.

The regional information which is gathered from typical images is used to label the corresponding regions in the FEM mesh. the 16 optical channels are simulated, as is the standard practice in human imaging studies. They are used to collect data using 16 sources and 16 detectors (giving 240 measurements) in a single plane geometry so that there are one source and one detector for each 16 optical channels. Sources are considered to be intensity modulated with a frequency of 100 MHz, and measurements consisting of the logarithmic modulation amplitude b^A and phase shift b^φ, so that the target data $b = (b^A, b^\varphi)$ is calculated with the FEM diffusion forward model. With this model, the object is discretized into 2880 non-overlapping triangular elements connected by 1501 nodes, defining a piecewise quadratic unstructured basis expansion. Both b^A and b^φ are then contaminated with 2% zero mean additive Gaussian random noise which is always considered in practical simulations. For a prior information, pixels of similar intensity as segmented MRI are assumed to represent the same material or tissue such as [12].

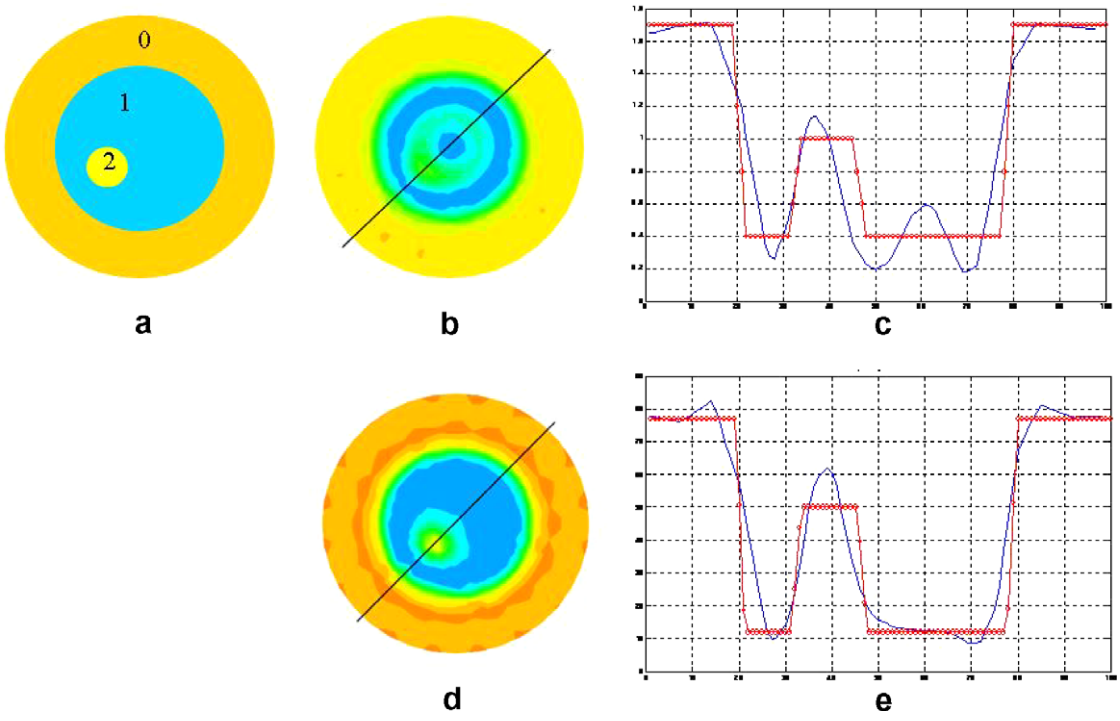

Fig. 2. (a) The simulated test phantom, (b) the reconstructed image with traditional Tikhonov regularization, and (c) its associated transect at 45°. (d) The reconstructed image with proposed preconditioning method, and (e) its associated transect at 45°.

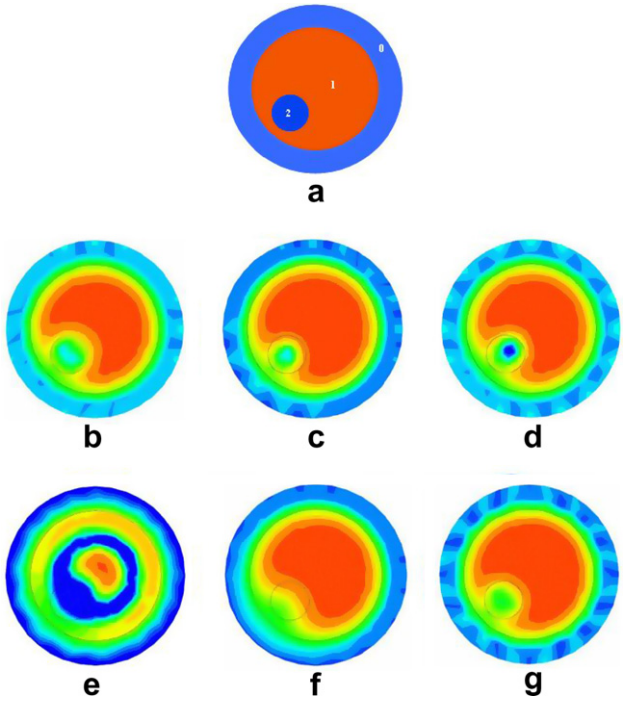

Fig. 3. (a) The simulated test phantom, (b) the reconstructed image with traditional MAP estimator, (c) the reconstructed image using proposed priorconditioners, (d) using RTLS to proposed priorconditioners without PCA, (e) the reconstructed image with Traditional TSVD for dimension reduction of proposed priorconditioning method, (f) using proposed PCA for dimension reduction of proposed priorconditioning method (g) using RTLS to proposed PCA method.

5.2. Phantom imaging: proposed method without PCA

Fig. 2b gives the reconstructed image of μ_a for Tikhonov regularization with regularization matrix L and regularization parameter α which are calculated as Section (3.2). By comparison, Fig. 2d gives the reconstructed image for the same phantom by proposed method in Eq. (14) which the preconditioners are constructed by assumption Gaussian measurement noise and access to typical measurement for data and anomaly. The proposed method with inherent regularization and MAP estimator which are applied as priorconditioner to linear systems, has flexible and less computational operations with CGLS and GMRES iterative algorithms. The associated transect of Fig. 2b at 45° that is plotted to Fig. 2c has the better quality for anomaly detection with respect to corresponding traditional Tikhonov regularization that is plotted to Fig. 2e. The reconstruction results to reconstruct Fig. 2d represent the 14th iteration of the GMRES method, taking 50 min on a 2.8 GHz pentium processor with 1 GB RAM wile the corresponding values by the traditional Tikhonov regularization method is 17 iteration and 62 min, respectively.

5.3. Phantom imaging: using RTLS and PCA

Fig. 3b gives the reconstructed image of μ_a when using traditional MAP estimator with sample based image and noise covariance matrixes as Section (3.3). By comparison, Fig. 3c gives the reconstructed image for the same phantom with proposed priorconditioners as previous section. The details of how RTLS criteria can improve the quality of image is shown in Fig. 3d. here, we have used RTLS to preconditioned DOT linear equation of type (14) with iterative methods. The better results of RTLS are shown when using whitened PCA model of type (20) by comparison of three latter images. Fig. 3e gives the reconstructed image when traditional TSVD model reduction is applied to proposed priorconditioners (14) without using RTLS, that is discussed in Section (2.2). in this image, the regions are faded. Fig. 3f gives the reconstructed image by proposed PCA method for dimension reduction of the covariance matrix which is constructed by accessing to typical measurement for data and anomaly, as (17). Comparison of Fig. 3e and f would demonstrates that dimension reduction with proposed PCA method has the superior quality in anomaly detection in respect to TSVD. The better results are shown when we used RTLS to Eq. (20) that is shown in Fig. 3g. Besides, necessity of changing the system matrix A and so calculating the SVD of it in any iteration, TSVD method has inherent computational complexity with respect to applying RTLS to proposed PCA method.

6. Conclusion

In this work, we formulated the linearized DOT problem within a preconditioning framework where the constructing of them is based on the *a priori* information extracted from a training set of typical and anomaly based images and their observation measurements. This help us to apply RTLS criteria for data and model matching which is used in our simulations. In another work, we used model reduction based on PCA to image covariance matrix to improve the reconstruction speed. The reconstructed results show that using preconditioners has the better results and improves the reconstructed image quality dramatically as compared to the Tikhonov regularization. In addition, using PCA to proposed method has the better image with respect to TSVD model reduction. The images show that applying RTLS can improve the reconstructed image specially when using PCA based model reduction. Quantitatively, priorconditioner results are close to the expected results.

References

[1] C. Guy, D. Fytche, An Introduction to the Principles of Medical Imaging, Imperial College Press, London, 2000.
[2] S.R. Arridge, J.P. Kaipio, V. Kolehmainen, M. Schweiger, E. Somersalo, T. Tarvainen, M. Vauhkonen, Inverse Probl. 22 (2006) 175.
[3] J. Selb, D. Joseph, Biomed. Opt. 11 (4) (2006).
[4] K. Ren, G. Bal, A.H. Hielscher, SIAM J. Sci. Comput. (2006) 1463.
[5] A. Cong, G. Wang, Opt. Express 13 (24) (2005) 9847.
[6] A.H. Hielscher, A.G. Yodh, IEEE Trans. Med. Imag. 20 (6) (2001) 2137.

[7] N. Cao, A. Nehorai, Opt. Express 15 (3) (2007) 896.
[8] A.P. Gibson, J.C. Hebden, S.R. Arridge, Phys. Med. Biol. 50 (2005) R1.
[9] M. Schweiger, S.R. Arridge, I. Nissil, Phys. Med. Biol. 50 (2005) 2365.
[10] R.A. Renaut, H. Guo, SIAM J. Matrix Anal. Appl. 26 (2) (2005) 457.
[11] H. Guo, R.A. Renaut, Iterative Method for Regularized Total Least Squares, Department of Mathematics and Statistics, Arizona State University, 2005.
[12] B. Brooksby, H. Dehghani, B.W. Pogue, IEEE J. Quantum Electron. 9 (2003) 199.
[13] B. Kanmani1, R.M. Vasu, Phys. Med. Biol. 50 (2005) 247.
[14] S.R. Arridge, M. Schweiger, Opt. Express 2 (1998) 213.
[15] A.H. Hielscher, A.D. Klose, K. Hanson, IEEE Trans. Med. Imag. 18 (3) (1999).
[16] M. Hanke, J.G. Nagy, R.J. Plemmons, in: L. Reichel, A. Ruttan, R.S. Varga (Eds.), Numerical Linear Algebra and Scientific Computing, de Gruyter, Berlin, 1993, p. 141.
[17] A.N. Tkhonov, V. Arsenin, Solution of Ill-posed Problems, Winston, Washington, DC, 1997.
[18] N. Bellomo, A. Bellouquid, Math. Models Meth. Appl. Sci. 14 (2004) 1683.
[19] H. Anton, C. Rorres, Linear Algebra: Applications Version, eighth ed., John Wiley and Sons, New York, 2000.
[20] C. Vogel, Computational Methods for Inverse Problems, Society for Industrial and Applied Mathematics, 1999.
[21] D.H. Brooks, G.F. Ahmad, R.S. MacLeod, IEEE Trans. Biomed. Eng. 46 (1999) 3.
[22] D. Calvetti, L. Reichel, A. Shuibi, Appl. Numer. Math. 54 (2005) 135.
[23] D.A. Benaron, J. Cereb. Blood Flow Metab. 20 (2000) 469.
[24] F. Gao, H.J. Zhao, Y. Yamada, Appl. Opt. 41 (2002) 778.
[25] A.R. Zirak, K. Mafinejad, in: Proceedings of the 8th ICEE Conference, Iran, 2000, p. 415.
[26] J. Kaipio, E. Somersalo, Statistical and Computational Inverse Problems, Springer, 2004.
[27] T. Huttunen, J. Kaipio, K. Hynynen, IEEE Trans. UFFC 50 (2003) 1486.
[28] M. Guven1, B. Yazici1, Phys. Med. Biol. 50 (2005) 2837.
[29] A.R. Zirak, M. Khademi, in: Optics and Photonics Conference, Shiraz University of Iran, 2006.