On iterative algorithms for quantitative photoacoustic tomography in the radiative transport regime

Chao Wang and Tie Zhou

School of Mathematical Sciences, Peking University, People's Republic of China

E-mail: chaowyww@pku.edu.cn and tzhou@math.pku.edu.cn

Received 8 April 2017, revised 14 August 2017
Accepted for publication 1 September 2017
Published 10 October 2017

Abstract
In this paper, we present a numerical reconstruction method for quantitative photoacoustic tomography (QPAT), based on the radiative transfer equation (RTE), which models light propagation more accurately than diffusion approximation (DA). We investigate the reconstruction of absorption coefficient and scattering coefficient of biological tissues. An improved fixed-point iterative method to retrieve the absorption coefficient, given the scattering coefficient, is proposed for its cheap computational cost; the convergence of this method is also proved. The Barzilai–Borwein (BB) method is applied to retrieve two coefficients simultaneously. Since the reconstruction of optical coefficients involves the solutions of original and adjoint RTEs in the framework of optimization, an efficient solver with high accuracy is developed from Gao and Zhao (2009 Transp. Theory Stat. Phys. 38 149–92). Simulation experiments illustrate that the improved fixed-point iterative method and the BB method are competitive methods for QPAT in the relevant cases.

Keywords: quantitative photoacoustic tomography, improved fixed-point iteration, radiative transfer equation, Barzilai–Borwein method

(Some figures may appear in colour only in the online journal)

1. Introduction

Photoacoustic tomography (PAT) is a developing medical imaging technique using non-ionizing waves. It combines ultrasound imaging and optical imaging [29], and so can achieve high contrast and high resolution simultaneously. Photoacoustic imaging is based on photoacoustic effect. A short pulse of electromagnetic energy propagates through the tissue rapidly at speed of higher magnitude than sound, after the tissue is illuminated. Then the electromagnetic
energy is absorbed by the tissue and converted to heat. An increase in tissue temperature quickly follows, which leads to spatial expansion, which in turn induces a pressure field. Acoustic waves generated by pressure distribution can then be measured by detectors placed at the surface. These can be used to extract physical information about the tissue inside. In medical imaging science, this technique is mostly applied to tomography imaging of skin, breast cancer screening and small animal imaging [29].

In the physical view, there are two inverse problems associated with two stages: the first stage of generating pressure after absorbing electromagnetic energy, and the second stage of initial pressure generating acoustic waves. The first problem—photoacoustic tomography (PAT)—is to reconstruct the initial pressure distribution from temporal measurements of photoacoustic waves, and the second one, called quantitative PAT (QPAT), is to estimate optical coefficients from initial pressure distribution. Although initial pressure reflects some information inside, it is the product of the absorption of energy, indirectly dependent on properties of the object. Only the intrinsic optical coefficients provide direct information. Essentially, spatially varying optical coefficients lead to spatially varying initial pressure distribution. Moreover, [22] mentions that initial pressure decreases quickly in tissue areas with small absorption coefficient, so that we cannot distinguish visually in the final image of initial pressure. To summarize, the original information in the investigated object cannot be obtained completely until the optical coefficients are recovered. Generally, estimated optical coefficients comprise absorption coefficient, scattering coefficient and Grüneisen coefficient, where absorption coefficient is usually of principal clinical interest [14]. Nowadays, algorithms on reconstruction of initial pressure have been developed maturely [1, 20, 21, 28, 29], so a lot of researchers are focusing on QPAT.

In terms of mathematical modeling of QPAT, light propagation inside the tissue can be modeled by radiative transfer equation (RTE) [18, 22, 25] or diffusion approximation (DA) [4, 12]. However, DA approximates to RTE only in the situation where light behaves diffusely, which is invalid for small objects or in regions close to the light source, so that the optical coefficients estimated from the DA model are not accurate. The reconstruction theories and some algorithms based on the DA model are presented in [5, 7, 13, 23, 24, 31]. The two models are compared numerically in [27], and the derivation from RTE model to DA model can be found in [10].

In terms of measurement, there are three main kinds of methodology: single illumination, multi-source illumination and multi-spectral illumination. It has been proved that single illumination can recover any one of absorption and Grüneisen coefficient stably, given scattering coefficient, and there is even an explicit analytic formula for non-scattering media [22]. Multi-source illumination can recover two of the three coefficients mentioned above uniquely and simultaneously [3, 18, 22, 26]. Furthermore, multi-spectral illumination can achieve unique reconstruction of three coefficients [2]. We focus on the RTE model with multi-source illumination and aim to investigate more efficient reconstruction method in this setting.

In terms of reconstruction algorithms, assuming the knowledge of scattering coefficient, fixed-point iteration is used to recover absorption coefficient in [19]. However, the requirement of close enough initial value, and the lack of theoretical results, are obstacles to its application. When the scattering coefficient is unknown, QPAT can be formulated as a nonlinear optimization problem [14, 18], and furthermore can be simplified to a linear optimization problem by Born linearization [22]. Optimization approaches such as Jacobian-based and gradient-based methods are investigated in [25], in which limited-memory Broyden–Fletcher–Goldfarb–Shanno (LBFGS), as a state of art gradient-based method, has been applied to the problem [14].

In the optimization framework, it is inevitable to solve original RTE in forward problem and adjoint RTE in the process of calculating gradients of objective functions. The finite element method, combined with streamline diffusion modification, is applied to solve stationary
RTE [18, 25, 30]. In this way, a large sparse linear system needs to be solved, which is still
difficult. Inspired by [15], we apply the discontinuous Galerkin (DG) method, combined with
multigrid methodology, to solve the two RTEs.

Our contributions in this paper are severalfold.

(i) We propose an improved fixed-point iterative algorithm, and prove its convergence for
a given scattering coefficient. Numerically, the improved fixed-point iteration converges
faster than the optimization-based Barzilai–Borwein (BB) method.

(ii) We apply the BB method to reconstruct the absorption coefficient and scattering coeffi-
cient simultaneously. To the best of our knowledge, this is the first time the BB method
has been applied to this problem.

(iii) We revise the numerical scheme for RTE proposed in [15], which fails to solve adjoint
RTE. We revise it to make it suitable for the adjoint RTE, and prove its convergence.

Thus, the RTE can finally be reduced to a sparse block diagonal linear system.

The rest of the paper is organized as follows. We introduce mathematical models of the
forward problem and inverse problem, as well as notational conventions, in section 2. Given
scattering coefficient, an improved fixed-point iterative algorithm and relevant proofs for
convergence are presented in section 3.1. The BB method for QPAT, and the deduction of the
gradient of objective function, are detailed in section 3.2. We discuss solvers for original RTE
and adjoint RTE in section 4.1. The simulation results are shown in section 4.2. The conclu-
sion comes in section 5.

2. Mathematical description

In the following, we consider the convex bounded object region Ω ∈ ℜ^n (n = 2, 3) with
Lipschitz boundary ∂Ω and angular space S^{n-1}. Describing boundary conditions, we divide
the boundary ∂Ω into inflow boundary and outflow boundary

\[ \Gamma_- := \{(x, \theta) \in \partial \Omega \times S^{n-1} : \nu(x) \cdot \theta < 0 \} , \]

\[ \Gamma_+ := \{(x, \theta) \in \partial \Omega \times S^{n-1} : \nu(x) \cdot \theta > 0 \} , \]

where ν(x) is the outward normal vector at the position x of the boundary ∂Ω.

2.1. QPAT model

We use stationary RTE to describe the light propagation in the form of

\[ [\theta \cdot \nabla + \mu_a(x) + \mu_s(x)] \phi(x, \theta) - \mu_s(x) \int_{S^{n-1}} f(\theta, \theta') \phi(x, \theta') d\theta' = q(x, \theta), \quad (x, \theta) \in \Omega \times S^{n-1}, \]

where θ and x are the direction and the position of interest respectively; \( \phi(x, \theta) \) is the density
of energy at the position x in direction \( \theta \); \( \mu_a(x) \) and \( \mu_s(x) \) are spatially varying absorption and
scattering coefficients respectively; and \( q(x, \theta) \) is the source term. Function \( f(\theta, \theta') \), describing
the probability of a photon traveling from direction \( \theta' \) to \( \theta \), is called the scattering phase
function, and is usually characterized by anisotropic factor g of the form

\[ f(\theta, \theta') = \begin{cases} 
\frac{1 - g^2}{2\pi(1 - g^2 - 2g\theta \cdot \theta')}, & n = 2, \\
\frac{1 - g^2}{4\pi(1 - g^2 - 2g\theta \cdot \theta')^{3/2}}, & n = 3,
\end{cases} \]
which is the well-known Henyey–Greenstein (H–G) scattering function. For the sake of simplicity, we define the scattering operator \( K \) by
\[
(K\phi)(x, \theta) := \int_{S_{n-1}} f(\theta, \theta')\phi(x, \theta') \, d\theta'.
\] (4)

We assume inflow boundary condition
\[
\phi(x, \theta) = q_b(x, \theta), \quad (x, \theta) \in \Gamma_-, \quad (x, \theta) \in \Gamma_-
\] (5)

that is, no photon enters \( \Omega \) except from the boundary source \( q_b(x, \theta) \). In QPAT, the source term \( q(x, \theta) \) is usually regarded as zero. Absorbed energy density is presented by
\[
h(x) := \mu_a(x)\Phi(x),
\] (6)

where
\[
\Phi(x) := \langle A\phi \rangle(x) = \int_{S_{n-1}} \phi(x, \theta) \, d\theta,
\] (7)
in which \( A \) is the accumulation operator over all directions.

Owing to absorbed energy, initial acoustic pressure distribution is generated according to the form
\[
p_0(x) := \gamma(x)h(x),
\]
where \( \gamma(x) \) is spatially varying Grüneisen coefficient. \( \gamma(x) \) is the efficiency of conversion of absorbed energy to pressure and is assumed to be known. Throughout this work, it is rescaled to be one, i.e. \( \gamma(x) = 1 \).

2.2. PAT model

On account of initial pressure distribution, acoustic waves propagate through object region \( \Omega \) and a series of temporal acoustic signals are detected by ultrasonic detectors at the surface \( \partial\Omega \). The behavior is described by
\[
\begin{cases}
\frac{1}{c(x)} \frac{\partial^2}{\partial t^2} p(x, t) - \Delta p(x, t) = 0, & (x, t) \in \mathbb{R}^n \times (0, T], \\
p(x, 0) = p_0(x), & x \in \Omega, \\
\frac{\partial p}{\partial n}(x, 0) = 0, & x \in \Omega,
\end{cases}
\] (8)

where \( c(x) \) is the acoustic speed, \( p(x, t) \) is the acoustic pressure at position \( x \) and time \( t \), and \( p_0(x) \) is the initial pressure distribution.

The goal of PAT is to recover optical coefficients \( \mu_a \) and \( \mu_s \), given temporal data \( p(x, t) \) ((\( x, t \)) \( \in \partial\Omega \times (0, T)] \). This is often accomplished in two stages: first one recovers \( p_0(x) \) from temporal data \( p(x, t) \) ((\( x, t \)) \( \in \partial\Omega \times (0, T)] \), and next one reconstructs \( \mu_a \) and \( \mu_s \) from initial pressure \( p_0 \) (or \( h \)), given illumination conditions. We assume the first stage has been done, and that the second stage QPAT is our concern. We assume tissue is illuminated for \( M \) times with known boundary sources \( q_{bm} \) (\( m = 0, 1, \ldots, M - 1 \)); then, the inverse problem QPAT is to determine \( \mu_a \) and \( \mu_s \) from data \( h_m \), that is \( h(x; \mu_a, \mu_s, q_{bm}) \) (\( m = 0, 1, \ldots, M - 1 \)).

2.3. Notation conventions

We denote the absorption coefficient and scattering coefficient in the \( i \)th iteration by \( \mu_a^i \) and \( \mu_s^i \), respectively, and corresponding solution of RTE and heat with \( m \)th illumination of boundary
source $q_{b_m}$ ($m = 0, 1, \ldots, M - 1$) is denoted by $\phi_{m}^0(x, \theta)$ and $h_{m}^0(x)$. The exact coefficients and data are denoted by $\mu_{a_i}^*, \mu_{s_i}^*, \phi_{m}^*$, and $h_{m}^*$.  

3. Reconstruction for optical coefficients

In this section, we propose an improved fixed-point iterative method to recover absorption coefficient given scattering coefficient and then apply Barzilai–Borwein (BB) method to simultaneously recover scattering coefficient as well. First, it is necessary to make several assumptions regarding the two optical coefficients.

**Assumption 1.** Assume optical coefficients $\mu_{a_i}$, $\mu_{s_i}$ and energy density $\phi_{m}(x, \theta; \mu_{a_i}, \mu_{s_i})$, which is the solution of RTE with $\mu_{a_i}$ and $\mu_{s_i}$, satisfy

(i) $\mu_{a_i}(x) \in D_{a_i} := \{ \mu_{a_i} \in C(\Omega) : 0 < \mu_{a_i}^0 \leq \mu_{a_i}(x) \leq \mu_{a_i}^{upper} \}$ for fixed $\mu_{a_i}^{upper}$;

(ii) $\mu_{s_i}(x) \in D_{s_i} := \{ \mu_{s_i} \in C(\Omega) : 0 < \mu_{s_i}^0 \leq \mu_{s_i}(x) \leq \mu_{s_i}^{upper} \}$ for fixed $\mu_{s_i}^{upper}$;

(iii) Boundary condition $q_{b_m}(m = 0, 1, \ldots, M - 1) \in L_{\infty}(\Gamma_{-}, |\theta| \cdot \nu)$;

(iv) Scattering kernel $f(\theta, \theta') \in L^1(S^{n-1} \times S^{n-1})$;

(v) If $\mu_{a_i}$ and $\mu_{s_i}$ satisfy (i) and (ii), there exist boundary sources $q_{b_m}(m = 0, 1, \ldots, M - 1)$ such that $\phi_{m}(x, \theta) \geq \epsilon > 0$ for some positive $\epsilon$.

3.1. The case of given scattering coefficient

Under assumption 1, we provide an effective fixed-point iterative method. Since one measurement can recover the absorption coefficient, subscripts $m'$ of symbols mentioned in section 2.3 are omitted for brevity. The algorithm is detailed as follows.

**Algorithm 1.** Improved fixed-point iteration.

**Input:** Given initialization $\mu_{a_i}^0$ mentioned in assumption 1, data $h^*$, scattering coefficient $\mu_{s_i}$, boundary source $q_{b_m}$, and tolerance $\epsilon$.

1: for $i = 0, 1, \ldots$ do

2: \hspace{10pt} Solve stationary RTE with absorption $\mu_{a_i}^i$ and scattering coefficients $\mu_{s_i}$ to obtain the solution $\phi_i$. Then let $h_i(x) = \mu_{a_i}^i(x) \langle A \phi_i \rangle(x)$;

3: \hspace{10pt} If $\|h_i - h^*\|_1 < \epsilon$, break with $\mu_{a_i} = \mu_{a_i}^i$; or implement next step;

4: \hspace{10pt} Calculate $\mu_{a_i}^{i+1}(x) = \frac{h_i(x)}{\langle A \phi_i \rangle(x)}$;

5: \hspace{10pt} Calculate $\mu_{a_i}^{i+1}(x) = \max\{\mu_{a_i}^i(x), \mu_{a_i}^{i+1}(x)\}$.

6: end for

**Remark.** The assumption 1 is important because $\phi$ is a denominator in the algorithm 1. In practice, we can also use $\phi + \epsilon$ instead with the small $\epsilon$ to avoid the situation where $\phi = 0$.

A fixed-point iterative method has been mentioned in many references [9, 11, 14, 19], which is usually unstable and sensitive to initial values. However, the modification of the fifth step in algorithm 1 can guarantee convergence for a given scattering coefficient.

Carrying out algorithm 1, we can obtain function sequences $\{\mu_{a_i}^{i+1}(x)\}_{i=0}^\infty$, $\{\mu_{a_i}^i(x)\}_{i=0}^\infty$, $\{h_i(x)\}_{i=0}^\infty$ and $\{\phi_i(x)\}_{i=0}^\infty$. We claim that the sequence $\|h_i - h^*\|_1$ converges to zero. The proof is presented as follows.
Firstly, we state a theorem on stationary RTE.

**Theorem 1 ([8]).** For RTE (2) satisfying assumption 1, if source \( q \geq 0 \) and boundary source \( q_b \geq 0 \), there exists a unique non-negative and continuous solution \( \phi(x, \theta) \).

From the physics perspective, the intuitive explanation behind theorem 1 is straightforward. For non-negative initial energy, the energy will not become negative while being transferred and absorbed. Lemma 2 reveals the monotonicity of \( \phi \) with respect to \( \mu_a \).

**Lemma 2.** Let \( \phi^1(x) \) and \( \phi^2(x) \) be the solutions of RTEs

\[
\begin{aligned}
\left\{ & \[ \theta \cdot \nabla + \mu_a(x) + \mu_s(x) \] \phi(x, \theta) - \mu_s(x)(K\phi)(x, \theta) = 0, \quad (x, \theta) \in \Omega \times \mathcal{S}^{n-1}, \\
& \phi(x, \theta) = q_b(x, \theta), \quad (x, \theta) \in \Gamma_-,
\end{aligned}
\]

(9)

with \( \mu_a \) being \( \mu^1_a \) and \( \mu^2_a \) respectively. Then \( \phi^1(x, \theta) \geq \phi^2(x, \theta) \) provided \( \mu^1_a(x) \leq \mu^2_a(x) \) \( \forall x \in \Omega \). Note that the superscripts of \( \mu^1_a \) and \( \mu^2_a \) are used to distinguish the different absorption coefficients.

**Proof.** We see that \( \phi := \phi^1 - \phi^2 \) solves the RTE

\[
\begin{aligned}
\left\{ & \[ \theta \cdot \nabla + \mu^2_a + \mu_s \] \phi(x, \theta) - \mu_s(K\phi)(x, \theta) = (\mu^2_a - \mu^1_a)\phi(x, \theta), \quad (x, \theta) \in \Omega \times \mathcal{S}^{n-1} \\
& \phi(x, \theta) = 0, \quad (x, \theta) \in \Gamma_-
\end{aligned}
\]

(10)

with absorption coefficient \( \mu^2_a \) and scattering coefficient \( \mu_s \). Since source term \((\mu^2_a - \mu^1_a)\phi\) and boundary source are both non-negative, we can derive

\[ \phi^1 \geq \phi^2 \]

(11)

using theorem 1. \( \square \)

Using lemma 2, we investigate the properties of function sequences from algorithm 1.

**Lemma 3.** For any \( x \in \Omega \), the sequence \( \{\mu^{i+1}_a(x)\}_{i=0}^{\infty} \) obtained in the algorithm 1 satisfies

\[ \mu^0_a(x) \leq \mu^1_a(x) \leq \cdots \leq \mu^i_a(x) \leq \cdots \leq \mu^*_a(x) \]

(12)

**Proof.** Given \( \mu^0_a(x) := \mu^0_a(x) \leq \mu^*_a(x) \), let us assume

\[ \mu^0_a(x) \leq \mu^1_a(x) \leq \cdots \leq \mu^i_a(x) \leq \mu^*_a(x) \]

(13)

Then it suffices to establish (13) for \( i + 1 \).

Obviously, \( \mu^i_a(x) \leq \mu^*_a(x) \) indicates \( \phi^i(x) \geq \phi^*(x) \). Since \( \mu^{i+1}_a(x) = \frac{h^*}{\mu^i_a(A\phi^*)} \) and \( \phi^i \geq \phi^* \), combining \( h^* = \mu^*_a(A\phi^*) \), we infer

\[ \mu^{i+1}_a(x) \leq \mu^*_a(x) \]

According to algorithm 1,

\[ \mu^i_a(x) \leq \mu^{i+1}_a(x) \leq \mu^*_a(x) \]

An easy induction completes the proof. \( \square \)

Monotonicity of \( \mu^i_a \) can easily deduce the monotonicity of \( \phi^i \).
Lemma 4. For any \((x, \theta) \in \Omega \times \mathcal{S}^{n-1}\), the sequence \(\{\phi^i(x, \theta)\}_{i=0}^{\infty}\) obtained in the algorithm 1 satisfies
\[
\phi^0(x, \theta) \geq \phi^1(x, \theta) \geq \cdots \geq \phi^i(x, \theta) \geq \cdots \geq \phi^*(x, \theta).
\]  

Proof. This is obvious, using lemmas 2 and 3.

We analyze the update process for \(\mu_a\) by introducing new definitions. For any integer \(i\), we divide region \(\Omega\) into two parts in the form of
\[
\begin{align*}
\Omega^i_+ &:= \{x \in \Omega : h^*(x) \geq h^i(x)\}, \\
\Omega^i_- &:= \{x \in \Omega : h^*(x) < h^i(x)\}.
\end{align*}
\]

Then we have
\[
\tilde{\mu}_a^{i+1} = \frac{h^*}{\Delta \phi} = \begin{cases} 
\mu_a^i, & x \in \Omega^i_+, \\
\mu_a^i, & x \in \Omega^i_-.
\end{cases}
\]

According to algorithm 1,
\[
\mu_a^{i+1} = \begin{cases} 
\tilde{\mu}_a^{i+1}(x), & x \in \Omega^i_+, \\
\mu_a^i, & x \in \Omega^i_-.
\end{cases}
\]

Actually, it is a process of keeping and updating from \(\mu_a^i\) to \(\mu_a^{i+1}\), keeping the value of \(\mu_a^i\) of \(\Omega^i_-\), updating the value of \(\Omega^i_+\).

Eventually, we prove the convergence of function sequence \(\{h^i(x)\}_{i=0}^{\infty}\).

Theorem 5. For any \(x \in \Omega\), the sequence \(\{h^i(x)\}_{i=0}^{\infty}\) obtained in the algorithm 1 satisfies
\[
\lim_{i \to \infty} \|h^* - h^i\|_1 = 0.
\]

Proof. We divide our proof into four claims.

(i) : Function sequences \(\{\mu_a^i(x)\}_{i=0}^{\infty}\), \(\{\phi^i(x, \theta)\}_{i=0}^{\infty}\) and \(\{h^i(x)\}_{i=0}^{\infty}\) converge pointwise.

Since monotonicity and boundedness of \(\{\mu_a^i(x)\}_{i=0}^{\infty}\) and \(\{\phi^i(x, \theta)\}_{i=0}^{\infty}\) for any \((x, \theta) \in \Omega \times \mathcal{S}^{n-1}\), it follows that they converge pointwise, with limits denoted by \(\mu_a^\infty(x)\) and \(\phi^\infty(x, \theta)\) respectively. Obviously, \(\{h^i(x)\}_{i=0}^{\infty}\) converges pointwise, and its limit is denoted by \(h^\infty(x)\). Again obviously, \(h_i(x) = \mu_a^i(x)(A \phi^i)(x)\).

(ii) : For any \(x \in \Omega\), if there is some \(i\) such that if \(x \in \Omega^i_+\), then \(x \in \Omega^i_+ (i > I)\).

It is obvious that \(\mu_a^i(x) = \mu_a^i h^i(x) \geq \mu_a^i(x)\). According to lemma 4, we can obtain
\[
h^i(x) = \mu_a^i (A \phi^i)(x) \leq \mu_a^i(x)(A \phi^i)(x) = h^i(x).
\]

An easy induction shows that \(x \in \Omega^i_+ (i > I)\).

(iii) : \(h_i(x) = h^\infty(x) (x \in \Omega^i_+), \|h^* - h^i\|_{L^1(\Omega^i_+)} \to 0 (i \to \infty)\).
Thanks to \( x \in \Omega_0^i (i > 0) \) for any \( x \in \Omega_0^i \), it follows that
\[
\begin{align*}
\mu_{a}^{i+1}(x) - \mu_a^i(x) &= \mu_{a}^{i+1}(x) - \mu_a^i(x) \\
&= \frac{h^*(x) - h'(x)}{A \phi(\theta)} \\
&\geq \frac{h^*(x) - h'(x)}{2\pi \phi(\theta)},
\end{align*}
\] (16)

where \( \phi = \sup_{(x, \theta) \in \Omega \times S^{n-1}} \phi(\theta,x, \theta) \). According to lemma 4,
\[
\phi(\theta,x, \theta) \leq \phi(\theta,x, \theta) \leq \phi(\theta),
\]
gives the last inequality in (16). Then
\[
\mu_{a}^{0}(x) - \mu_a^0(x) = \sum_{i=0}^{\infty} (\mu_{a}^{i+1}(x) - \mu_a^i(x)) \geq \frac{1}{2\pi \phi(\theta)} \sum_{i=0}^{\infty} (h^*(x) - h'(x)).
\]

Owing to \( 0 \leq \mu_{a}^{0} - \mu_a^0 < \infty \), we can obtain that
\[
h^*(x) - h'(x) \rightarrow 0 \ (\forall x \in \Omega_0^i),
\] (17)

that is, \( h^*(x) - h'(x) \) converges to 0 pointwise. So \( \mu_{a}^{0}(x) = h^*(x) \ (x \in \Omega_0^i) \). Integrating (16) with respect \( x \) over \( \Omega_0^i \) gives
\[
\| \mu_{a}^{i+1} - \mu_a^i \|_{L^1(\Omega_0^i)} \geq \frac{1}{2\pi \phi(\theta)} \| h^* - h' \|_{L^1(\Omega_0^i)}.
\]

Since
\[
\| \mu_{a}^{0} - \mu_a^0 \|_{L^1(\Omega_0^i)} = \int_{\Omega_0^i} (\mu_{a}^{0}(x) - \mu_a^0(x)) \, dx
\]
\[
= \int_{\Omega_0^i} \sum_{i=0}^{\infty} (\mu_{a}^{i+1}(x) - \mu_a^i(x)) \, dx
\]
\[
= \sum_{i=0}^{\infty} \int_{\Omega_0^i} (\mu_{a}^{i+1}(x) - \mu_a^i(x)) \, dx
\]
\[
\geq \frac{1}{2\pi \phi(\theta)} \sum_{i=0}^{\infty} \int_{\Omega_0^i} (h^*(x) - h'(x)) \, dx
\]
\[
= \frac{1}{2\pi \phi(\theta)} \sum_{i=0}^{\infty} \| h^* - h' \|_{L^1(\Omega_0^i)}
\]

and \( 0 \leq \| \mu_{a}^{0} - \mu_a^0 \|_{L^1(\Omega_0^i)} < \infty \), directly,
\[
\| h^* - h' \|_{L^1(\Omega_0^i)} \rightarrow 0 \ (i \rightarrow \infty).
\] (18)
(iv) \( \| h^* - h' \|_{L^1(\Omega_0^-)} \to 0 \ (i \to \infty) \).

We divide region \( \Omega_0^- \) into two parts of form
\[
\begin{align*}
\Omega_0^{01} &: = \{ x \in \Omega_0^- : \text{Exist some } i \text{ such that } x \in \Omega_i^+ \}, \\
\Omega_0^{02} &: = \{ x \in \Omega_0^- : \text{For any } i, x \in \Omega_i^+ \}.
\end{align*}
\]

Examining the proof of (iii) and the definition of \( \Omega_i^+ \), it can easily be seen that
\[
h(x) \begin{cases} = h^*(x), & x \in \Omega_0^{01}, \\
\geq h^*(x), & x \in \Omega_0^{02}, \end{cases}
\]
and
\[ \| h' - h^* \|_{L^1(\Omega_0^+)} \to 0. \] (20)

From (19), it is clear that
\[ \| h \|_{L^1(\Omega_0^+)} \geq \| h^* \|_{L^1(\Omega_0^+)}. \] (21)

Considering RTEs with absorption coefficient \( \mu_i^a \) and \( \mu_s^a \), it is easy to obtain
\[ \theta \cdot \nabla (\phi_i - \phi^*) + \mu_i^a \phi_i - \mu_s^a \phi^* = -(\mu J - \mu_K)(\phi_i - \phi^*). \] (22)

Since phase function \( f(\theta, \theta') \) satisfies
\[ \int_{S^{n-1}} f(\theta, \theta') \, d\theta = 1, \]
it follows that
\[
\int_{S^{n-1}} \mu_i(x)(K \phi)(x, \theta) \, d\theta = \int_{S^{n-1}} \mu_i(x) \int_{S^{n-1}} f(\theta, \theta') \phi(x, \theta') \, d\theta' \, d\theta \\
= \int_{S^{n-1}} \mu_i(x) \phi(x, \theta') \int_{S^{n-1}} f(\theta, \theta') \, d\theta' \\
= \int_{S^{n-1}} \mu_i(x) \phi(x, \theta) \, d\theta.
\]

Integrating (22) with respect to \( \theta \) gives
\[ h^i - h^* + \int_{S^{n-1}} \theta \cdot \nabla (\phi(x, \theta) - \phi^*(x, \theta)) \, d\theta = 0. \] (23)

Integrating (23) with respect to \( x \) over \( \Omega \) gives
\[ \| h^i \|_1 - \| h^* \|_1 + \int_{\Omega} \int_{S^{n-1}} \theta \cdot \nabla (\phi(x, \theta) - \phi^*(x, \theta)) \, d\theta \, dx = 0. \] (24)
Applying Green’s formula to the second term of (24), we have
\[
\int_{\Omega} \oint_{S} n^{-1} \theta \cdot \nabla (\phi'(x, \theta) - \phi^*(x, \theta)) \, d\theta \, dx
\]
\[
= \oint_{S}^{-1} \int_{\Omega} \theta \cdot \nabla (\phi'(x, \theta) - \phi^*(x, \theta)) \, dx \, d\theta
\]
\[
= \oint_{S}^{-1} \int_{\Gamma} (\theta \cdot \nu) (\phi'(x, \theta) - \phi^*(x, \theta)) \, dx \, d\theta
\]
\[
= \| \phi'(x, \theta) - \phi^*(x, \theta) \|_{L^1(\Gamma)}
\geq 0.
\]

Apparently, we have
\[\| h' \|_1 \leq \| h^* \|_1.\]

Naturally, when \( i \) is sufficiently large, we have
\[
\| h^* \|_1 - \| h' \|_1 = \int_{\Omega} (h^*(x) - h'(x)) \, dx
\]
\[
= \int_{\Omega} (h^*(x) - \bar{h}'(x)) \, dx - \int_{\Omega} (h'(x) - h^*(x)) \, dx
\]
\[
= \| h^* - h' \|_{L^1(\Omega)} - \| h' - h^* \|_{L^1(\Omega)}.
\]

Hence, when \( i \) tends to infinity, we have
\[
0 \leq \lim_{i \to \infty} \| h^* - h' \|_{L^1(\Omega)} - \| h^* - h' \|_{L^1(\Omega)} = 0 - \lim_{i \to \infty} \| h^* - h' \|_{L^1(\Omega)} \leq 0.
\]

Thus
\[
\lim_{i \to \infty} \| h^* - h' \|_{L^1(\Omega)} = 0.
\]

This completes the proof of (iv).

Therefore,
\[
\lim_{i \to \infty} \| h^* - h' \|_1 = 0.
\]

Given the scattering coefficient, by algorithm 1, we can get monotonically increasing sequence \( \mu_a'(x) \) and monotonically decreasing sequence \( \phi'(x, \theta) \), which converge pointwise to \( \mu_a \leq \mu^* \) and \( \phi \geq \phi^* \) respectively. Sequence \( h'(x) \) converges to exact function \( h^*(x) \) in the \( L^1 \)-norm.

### 3.2. Reconstruction of \( \mu_a \) and \( \mu_s \) simultaneously

In practice, the scattering coefficient is also unknown, so the improved fixed-point iteration method is not applicable in the case with two unknown coefficients. It is necessary to establish
a more general method to recover two coefficients simultaneously. From now on, we follow optimization approach to estimate \( \mu_a \) and \( \mu_s \). First, we define error functional

\[
\mathcal{F}(\mu_a, \mu_s) := \sum_{m=0}^{M-1} \frac{1}{2} \| \log(h_m(\mu_a, \mu_s)) - \log(h^*_m) \|_2^2
\]

(26)
to measure the distance between measurement data \( h_m^* \) and estimated data \( h_m \), where \( h_m(x; \mu_a, \mu_s) \) equals \( \mu_a(A\phi_m)(x; \mu_a, \mu_s) \) for estimated \( \mu_a \) and \( \mu_s \). Then the reconstruction of QPAT can be reformulated as

\[
\min_{\mu_a, \mu_s} \mathcal{F}(\mu_a, \mu_s).
\]

(27)

**Remark.** We replace \( \| h_m(\mu_a, \mu_s) - h_m^* \|_2^2 / 2 \) by \( \| \log(h_m(\mu_a, \mu_s)) - \log(h_m^*) \|_2^2 / 2 \) as data-fidelity term, since the latter boosts the convergence of the minimization method according to \[27\]. We discuss its advantage and properties in appendix C.

**Remark.** According to inverse problem theory, regularization is useful for handling ill-conditioned problems. Usually, in image science, ill-conditionedness results in the edges of images blurring, and in amplification of noise. Even though QPAT is a typical nonlinear problem, as pointed out in \[2, 22\], it is relatively stable with the help of multiple measurements. We focus on the minimization of data-fidelity term for the reason that multiple measurements alleviate the ill-posedness numerically. The algorithms on the minimization of objective function with regularization term, such as some a priori information, can be easily derived by our framework. We do not explore this in this paper.

Using the definition of Fréchet derivative \( \nabla \mathcal{F} \) of \( \mathcal{F} \) in feasible direction \((h_{\mu_a}, h_{\mu_s}) \in \mathcal{D}_a \times \mathcal{D}_s\), we can write

\[
\mathcal{F}'(\mu_a, \mu_s)(h_{\mu_a}, h_{\mu_s}) := \langle \nabla \mathcal{F}, (h_{\mu_a}, h_{\mu_s}) \rangle_{L^2(\Omega)}.
\]

(28)
The solution of RTE with coefficients \( \mu_a \) and \( \mu_s \) can be regarded as a function with respect to \( \mu_a \) and \( \mu_s \); that is, \( \phi = \phi(x, \theta; \mu_a, \mu_s) \). Fortunately, the directional derivative of \( \phi(x, \theta; \mu_a, \mu_s) \) with respect to \( \mu_a \) and \( \mu_s \) in any feasible direction exists according to \[18\]. Furthermore, the directional derivative of data-fidelity with respect to \( \mu_a \) and \( \mu_s \) can be expressed analytically by \( \mu_a, \phi(x, \theta; \mu_a, \mu_s) \) and its solution of adjoint RTE. This concludes in proposition 6.

Similarly to the proof of proposition 3.3 in \[18\], we can express the gradient of objective functional as follows.

**Proposition 6.** For any pairs \((\mu_a, \mu_s) \in \mathcal{D}_a \times \mathcal{D}_s\) and feasible direction \((h_{\mu_a}, h_{\mu_s}) \in \mathcal{D}_a \times \mathcal{D}_s\), we have

\[
\mathcal{F}'(\mu_a, \mu_s)(h_{\mu_a}, h_{\mu_s}) = \sum_{m=0}^{M-1} \frac{\log(\mu_a A\phi_m)}{\mu_a} - A(\phi_m \phi^*_m, h_{\mu_s}),
\]

\[
+ \sum_{m=0}^{M-1} (-\phi^*_m + (K\phi_m)\phi^*_m, h_{\mu_s}),
\]

(29)

where \( \phi^*_m \) solves following adjoint RTE

\[
\left\{ \begin{array}{l}
(-\theta \cdot \nabla_x + (\mu_a + \mu_s - \mu_s K))\phi^*_m = A^* \left( \frac{\log(\mu_a A\phi_m) - \log(h^*_m)}{A\phi_m} \right), \\
\phi^*_m|_{\Gamma^+} = 0.
\end{array} \right.
\]

(30)
Notice that $A^*$ is the adjoint operator of $A$ and

$$(A^*f)(x, \theta) = f(x), \quad \forall f \in L^2(\Omega).$$

**Proof.** Following the proof in [18] (proposition 3.3), we have

$$F'(\mu_a, \mu_s)(h_{\mu_a}, h_{\mu_s}) = M^{-1} \sum_{m=0}^{M-1} \left( \frac{\log(\mu_aA\phi_m(\cdot; \mu_a, \mu_s)) - \log(h_m^{\mu_a})}{\mu_a} \phi_m(\cdot; \mu_a, \mu_s) - A^* \frac{\log(\mu_aA\phi_m(\cdot; \mu_a, \mu_s)) - \log(h_m^{\mu_a})}{\phi_m} \phi_m(\cdot; \mu_a, \mu_s), h_{\mu_s} \right).$$

(31)

Notice that $\phi_m'(\cdot; \mu_a, \mu_s)(h_{\mu_a}, h_{\mu_s})$ is the directional derivative of $\phi_m(\cdot; \mu_a, \mu_s)$ with respect to $(\mu_a, \mu_s)$ in direction $(h_{\mu_a}, h_{\mu_s})$.

This completes the proof. □

After deducing the gradient of error functional (26), all we need to do is to find the appropriate step size in the negative gradient direction to decrease the functional. Many optimization methods can achieve this goal, including steepest descent, quasi-Newton, and so on. However, these gradient-based methods usually involve the linesearch process. A linesearch step needs to solve original RTE or adjoint RTE several times. Since solving RTE dominates the computational cost, linesearch is computationally intensive. To mitigate the heavy computational cost, the well-known BB gradient method is applied to compute stepsize. It is derived from a two-point approximation to the scant equation underlying quasi-Newton methods [6, 17]. And it is R-supralinearly convergent in the two-dimensional quadratic case [6]. Without loss of generality, we denote $\mu_a$ or $\mu_s$ by $\mu$, then the update formula at kth step is

$$\mu^{k+1} = \mu^k - \alpha_k \nabla F_k.$$

(32)

Generally, there are two choices about the stepsize $\alpha_k$:

$$\alpha_{k1} = \frac{s^T_k y_k}{\|y_k\|^2},$$

and

$$\alpha_{k2} = \frac{||s_k||^2}{s_k^T y_k},$$

where $s_k = \mu^k - \mu^{k-1}$ and $y_k = \nabla F_k - \nabla F_{k-1}$, and $F_k$ is the iterative sequence of error functional $F$, see (26). BB algorithm is detailed in algorithm 2.
Algorithm 2. BB method reconstruction.

Input: Given initialization \( \mu_0^a, \mu_0^s \), data \( h_m \), boundary source \( q_{b,m} \) \( (m = 0, 1, \ldots, M) \), \( \epsilon_1, \epsilon_2, \epsilon_3 > 0 \), maximum number of iterations \( N \), \( \text{flag}_a = 1 \) and \( \text{flag}_s = 1 \).

1: for \( i = 0, 1, \ldots, N \) do
2: \hspace{1em} If \( \text{flag}_a = 0 \) and \( \text{flag}_s = 0 \), end up with \( \mu_a = \mu_a^i \) and \( \mu_s = \mu_s^i \);
3: \hspace{1em} Solve stationary RTEs (2) with absorption and scattering coefficients \( \mu_a^i \) and \( \mu_s^i \) respectively to obtain the solution \( \phi_a^i \). Then let \( h_m^i(x) = \mu_a^i(x)(A \phi_a^i)(x) \);
4: \hspace{1em} If \( F_i < \epsilon_1 \), end up with \( \mu_a = \mu_a^i \) and \( \mu_s = \mu_s^i \);
5: \hspace{1em} Solve adjoint RTEs (30);
6: \hspace{1em} Calculate the gradient of \( F_i : \nabla_{(\mu_a, \mu_s)} F_i \) and \( \nabla_{\mu_a} F_i \) with respect to \( \mu_a \) and \( \mu_s \). If \( \| \nabla_{\mu_a} F_i \| \leq \epsilon_2 \) and/or \( \| \nabla_{\mu_s} F_i \| \leq \epsilon_3 \), let \( \text{flag}_a = 0 \) and/or \( \text{flag}_s = 0 \);
7: \hspace{1em} If \( i = 0 \) or 1, updating \( \mu_a \) and/or \( \mu_s \) in negative gradient direction with small step size such that \( F_i \) decreases; otherwise, if \( \text{flag}_a = 1 \) and/or \( \text{flag}_s = 1 \), update \( \mu_a \) and/or \( \mu_s \) by BB stepsize;
8: end for

4. Numerical simulations

The reconstructions of absorption and scattering coefficients are investigated with simulations in two cases: the reconstruction of \( \mu_a \) given \( \mu_s \) and the reconstruction of \( \mu_a \) and \( \mu_s \) simultaneously.

4.1. Solver for RTE

We consider the numerical solver for RTE in 2D; this can be extended to 3D with little effort. The finite element method, combined with streamline diffusion modification, is applied to solve stationary RTE [18, 25, 30], where \( P_1 \) Lagrangian elements in spatial and angular space are used. In this way, a large sparse linear system needs to be solved—which is still difficult. By improving the algorithm proposed in [15], we can solve the original RTE as well as adjoint RTE on 2D and 3D unstructured meshes by a DG method combined with multigrid methodology, which reduces the problem to a sparse block diagonal linear system.

We divide angular space into \( P \) equal intervals, and the corresponding directions and angles are denoted by \( \theta_0, \theta_1, \ldots, \theta_{p-1} \) and \( \beta_0, \beta_1, \ldots, \beta_{p-1} \) with interval \( \Delta \beta \). The spatial domain is discretized into unstructured triangular meshes. Suppose spatial domain is divided into \( N \) triangles, and each triangle contains \( n_d \) nodes, where \( n_d = 3 \) for 2D spatial domain. Lagrangian elements and piecewise linear DG elements are used to discretize RTE, that is, numerical discrete scheme of RTE is

\[
\phi(x, \theta) = \sum_{k=0}^{p-1} \sum_{i=0}^{N-1} \sum_{j=0}^{n_d-1} \phi_{i,j,k} \varphi_{ij}(x) L_k(\theta),
\]

(33)

where \( L_k(\theta) \) is the piecewise linear basis function in angular space which takes value of 1 in direction \( \theta_k \) and 0 in other directions, \( \varphi_{ij} \) is the linear basis function in spatial domain in the direction \( \theta_i \), which takes value of 1 in the \( k \)th nodes of \( i \)th triangle and 0 in other nodes and triangles, and \( \phi_{i,j,k} \) is the value of \( \phi(x, \theta) \) in direction \( \theta_i \) in the \( j \)th node of the \( k \)th element. This spatial basis function can approximate discontinuous solutions, thus being more suitable for actual situations, such as some edges of inclusions in object regions. On account of the scattering term in RTE, angular Gauss–Seidel iteration is applied in [15]; it iteratively solves RTE in every fixed direction in the form of
\[ \theta_k \cdot \nabla \phi_k + (\mu_a + \mu_s) \phi_k = \mu_s \sum_{k'=0}^{P-1} \omega_{kk'} \phi_{k'} + q_k, 0 \leq k \leq P - 1, \quad (34) \]

where \( \omega_{kk'} = \frac{\omega_{kk'}}{\sum_{k'} \omega_{kk'}} \) with \( \omega_{kk'}^0 = f(\theta_k, \theta_{k'}) \). Obviously, equation (34) can be solved by a number of numerical methods; we use the DG method.

A multigrid scheme is applied to solve (34), and it is reduced to solve a sparse block diagonal system. Multiplying test function \( \varphi_{ij} \) in both sides of (34), integrating it over \( i \)th triangle \( \tau_i \) with respect to \( x \), we have

\[
- \int_{\tau_i} \phi_k (\theta_k \cdot \nabla \varphi_{ij}) \, dx + \int_{\Gamma_+} \phi_k \varphi_{ij} (\theta_k \cdot \nu) \, dS + \int_{\tau_i} (\mu_a + \mu_s - \mu_s \omega_{kk}) \phi_k \varphi_{ij} \, dx
\]

\[
= - \int_{\Gamma_-} \hat{\phi}_k \varphi_{ij} (\theta_k \cdot \nu) \, dS + \int_{\tau_i} (\mu_s \sum_{k' \neq k} \omega_{kk'} \phi_{k'}) \varphi_{ij} \, dx + \int_{\tau_i} q_k \varphi_{ij} \, dx \quad (35)
\]

from Green’s formula, where subscripts \( i \) and \( j \) of \( \phi_{ij} \) are omitted. In (35), \( \hat{\phi} \) is the value of neighboring element in the inflow direction, which is the product of the upwind scheme. The specific explanation of why the value inflow direction is used is as follows.

For an intuitive explanation, we use rectangular mesh to illustrate in figure 1. For rectangular mesh, in direction \( \theta_k \) and node \( x \), RTE can be discretized into

\[
\cos \beta_k \frac{\partial \phi_k}{\partial x} + \sin \beta_k \frac{\partial \phi_k}{\partial y} + (\mu_a + \mu_s) \phi_k = \mu_s \sum_{k'=0}^{P-1} \omega_{kk'} \phi_{k'} + q_k, k = 0, 1, \ldots, P - 1, \quad (36)
\]

where \( \phi_k := \phi(x, \theta_k) \). Backward difference is used to approximate one order derivative. Then when \( 0 \leq \beta_k < \frac{\pi}{2}, \) its upwind scheme is

\[
(a + b + \mu_a + \mu_s - \mu_s \omega_{kk}) \phi_{i,j,k} - (a \phi_{i-1,j,k} + b \phi_{i,j-1,k}) - \mu_s \sum_{k' \neq k} \omega_{kk'} \phi_{i,j,k'} = q_{i,j,k}, \quad (37)
\]

where

\[ a = \frac{\cos \beta_k}{\Delta x} \geq 0, \quad b = \frac{\sin \beta_k}{\Delta y} \geq 0. \]

Obviously, the scheme converges because the equation (37) is diagonally dominant, and Gauss-Seidel scheme will accelerate the convergence. In fact, from (37), \( \phi_{i,j,k} \) is updated using
\[
\phi_{i,j,k} = \frac{a\phi_{i-1,j,k} + b\phi_{i-1,k} + \mu_s \sum_{k' \neq k} \omega_{kk'} \phi_{i,j,k'} + q_{i,j,k}}{a + b + \mu_a + \mu_s - \mu_s \omega_{kk}}.
\] (38)

Apparently, \(\phi_{i,j,k}\) and \(\phi_{i,0,k}\) are known as inflow boundary conditions. Therefore, (38) is an explicit scheme which updates \(\phi_{i,j,k}, \phi_{i,1,k}, \phi_{2,j,k}, \phi_{i,2,k}, \cdots\) \((i,j,k = 1, 2, \cdots)\) successively. We can see that (37) and (35) both use inflow information to update outflow information. This may be a kind of intuitive explanation. Indeed, the convergence of (35) for vacuum boundary condition is proved in [16]—see appendix A.

However, the above numerical scheme fails to solve the adjoint RTE with the same updating order—that is, to use inflow information to update outflow information. We provide a reverse updating order for elements to guarantee the convergence of solver for adjoint RTE. The solver for adjoint RTE is detailed as follows.

For adjoint RTE (30), the corresponding discrete scheme in direction \(\theta_k\) for rectangular mesh is

\[
- \cos \beta_k \frac{\partial \phi_k}{\partial x} - \sin \beta_k \frac{\partial \phi_k}{\partial y} + (\mu_a + \mu_s) \phi_k = \mu_s \sum_{k' = 0}^{P-1} \omega_{kk'} \phi_{k'} + \phi_k, \quad k = 0, 1, \ldots, P - 1.
\] (39)

In contrast, forward difference is used to approximate the first order derivative; when \(0 \leq \beta_k < \frac{\pi}{2}\), its upwind scheme is

\[
(a + b + \mu_a + \mu_s - \mu_s \omega_{kk}) \phi_{i,j,k} - (a\phi_{i+1,j,k} + b \phi_{i+1,k}) - \mu_s \sum_{k' \neq k} \omega_{kk'} \phi_{i,j,k'} = q_{i,j,k}.
\] (40)

where

\[a = \frac{\cos \beta_k}{\Delta x} \geq 0, \quad b = \frac{\sin \beta_k}{\Delta y} \geq 0.\]

Then \(\phi_{i,j,k}\) is updated by

\[
\phi_{i,j,k} = \frac{q_{i,j,k} + a\phi_{i+1,j,k} + b \phi_{i+1,k} + \mu_s \sum_{k' \neq k} \omega_{kk'} \phi_{i,j,k'}}{a + b + \mu_a + \mu_s - \mu_s \omega_{kk}}.
\] (41)

The use of forward difference makes (40) diagonally dominant again. Besides, outflow information is used to update inflow information in (41). Heuristically, for unstructured mesh, we have

\[
\int_{\Gamma_i} \phi_k \frac{\partial \varphi_{ij}}{\partial x} \, dx - \int_{\Gamma_i} \phi_k \varphi_{ij} \frac{\partial \theta_k}{\partial x} \, dS + \int_{\Gamma_i} (\mu_a + \mu_s - \mu_s \omega_{kk}) \phi_k \varphi_{ij} \, dx
\]

\[
= \int_{\Gamma_i} \phi_k \varphi_{ij} \frac{\partial \theta_k}{\partial x} \, dS + \int_{\Gamma_i} (\mu_s \sum_{k' \neq k} \omega_{kk'} \phi_{k'}) \varphi_{ij} \, dx + \int_{\Gamma_i} \phi_k \varphi_{ij} \, dx.
\] (42)

where \(\hat{\phi}\) is the value of the neighboring element in the outflow direction which is used to update inflow information. Using similar discussion, the convergence of (42) is proved in appendix B. Furthermore, we apply multigrid methodology to solve the RTE, to accelerate convergence.

From the above discussion, no matter which mesh is used, the key of convergence is to update outflow information using inflow information for the original RTE (2) and update the latter using the former for the adjoint RTE (30). The convergence in rectangular mesh is obvious. As for triangular mesh, the update scheme results from variation analysis, so the
convergence proofs are obtained by discussing corresponding variations—see appendices A and B.

Therefore, for the original RTE (2), there are three layers of loops. We apply the multigrid scheme in the outermost loop. Since it involves two coordinate systems, i.e. angular-coordinate and spatial-coordinate, there are variable updating schemes in terms of the iteration order, such as the angle-prior and space-prior. The second-layer loop is about direction—that is, the corresponding spatial equation (34) is solved in turn for each direction $\theta_k (k = 0, 1, \ldots, P - 1)$.

In the third-layer loop, for each element in the direction $\theta_k$, the value of $\phi_{i,j,k}$ is updated iteratively through solving a $3 \times 3$ linear system (35). Note that the updating order is in the following order: consider the each element successively from the boundary along the direction of $\theta_k$. We refer the interested reader to [15] for its algorithm details, and to [16] for its theory. In contrast, for the adjoint RTE, in the third-layer loop, inspired by (40) we propose to consider each element successively from the boundary along the direction of $-\theta_k$, which just reverses the updating order of the original RTE. The linear system (42) needs to be solved. The corresponding pseudo-code is presented in algorithm 3.

\begin{algorithm}
\caption{Original and adjoint RTE solver.}
1: for each loop of multigrid iteration do
2: \hspace{1em} for each direction $\theta_k (k = 0, 1, \ldots, P - 1)$ do
3: \hspace{2em} for each element $\tau_i$ do
4: \hspace{3em} Updating $\phi_{i,j,k} (j = 0, 1, 2)$ by solving a $3 \times 3$ linear system (35) for original RTE or (42) for adjoint RTE. For original RTE, the updating order is from boundary along $\theta_k$ through each elements to other side of region. For adjoint RTE, the updating order is reverse.
5: \hspace{2em} end for
6: end for
7: end for
\end{algorithm}

4.2. Numerical results

In this section, we present some numerical results to demonstrate the numerical performance of our improved fixed-point iteration and BB method. For the sake of simplicity, only 2D reconstruction is investigated. The anisotropic factor $g$ equals 0.9. Applying RTE solver described in section 4.1, we can get discrete energy density $H$ of the same length as the mesh. In order to explore the stability to noise, noisy data are generated by

$$\tilde{H} = H(1 + \epsilon N),$$

where $N$ is a random vector that follows the normal distribution with mean 0 and variance 1. We consider four object regions:

(i) Rectangle $\Omega_0 = [−20, 20] \times [−20, 20]$ and four inclusions $\Omega_1 = \{(x, y) \in \Omega_0 : (x + 10)^2 + (y − 10)^2 = 5^2\}$, $\Omega_2 = \{(x, y) \in \Omega_2 : (x − 10)^2 + (y − 10)^2 = 3^2\}$, $\Omega_3 = [−17, −5] \times [−17, −5]$; and $\Omega_4 = [5, 17] \times [−17, −5]$;

(ii) Circle $\Omega_0$ with center $(0, 0)$ and radius 20;

(iii) Circle $\Omega_0$ with center $(0, 0)$ and radius 20, and four inclusions $\Omega_1 = [−12, −8] \times [−12, 12]$, $\Omega_2 = [−8, −2] \times [−12, 12]$, $\Omega_3 = [−2, 12] \times [6, 12]$, and $\Omega_4 = [−2, 12] \times [−12, 6]$;

(iv) Circle $\Omega_0$ with center $(0, 0)$ and radius 20, and three inclusions $\Omega_1 = \{(x, y) \in \Omega_0 : (y − 3)^2 + (x − 7)^2 = 1\}$, $\Omega_2 = [−14, −4] \times [−10, 8]$, and $\Omega_3 = \{(x, y) \in \Omega_0 : \left(\frac{y^2}{3} + \frac{x^2}{8} − 8.4\right)^2 + \left(\frac{y^2}{3} + \frac{x^2}{8} + 8\right)^2 = 1\}$. 


Their respective absorption and scattering coefficients are

(i) \( \mu_a = 0.02 \) in \( \Omega_1 \) and \( \Omega_4 \) with background \( 0.01 \); \( \mu_s = 3 \) in \( \Omega_2 \) and \( \Omega_3 \) with background \( 1 \),

(ii) \( \mu_a(x, y) = 0.02 + 0.01 \sin\left(\frac{\pi x}{8}\right) \) and \( \mu_s(x, y) = 2 + \sin\left(\frac{\pi y}{8}\right) \),

(iii) \( \mu_a = 0.03 \) in \( \Omega_1 \), \( 0.02 \) in \( \Omega_2 \), \( 0.04 \) in \( \Omega_3 \), and \( 0.015 \) in \( \Omega_4 \) with background \( 0.01 \); \( \mu_s = 2.5 \) in \( \Omega_1 \), \( 1.5 \) in \( \Omega_2 \), \( 3 \) in \( \Omega_3 \), and \( 2 \) in \( \Omega_4 \) with background \( 1 \),

(iv) \( \mu_a = 0.015 \) in \( (\Omega_1 \cup \Omega_2) \setminus (\Omega_1 \cap \Omega_2) \cup \Omega_3 \) and \( 0.03 \) in \( \Omega_1 \cap \Omega_2 \) with background \( 0.01 \); \( \mu_s = 3 \) in \( \Omega_1 \cup \Omega_3 \) with background \( 1 \).

These regions are depicted in figure 2, where the optical coefficients of the first, third and fourth templates are piecewise constant, and the second one is smooth in object region \( \Omega_0 \). These templates contain both continuous and discontinuous borders, whose corners are both sharp and smooth. To avoid inverse crime, we generate data by solving RTE in finer unstructured mesh than inverse problem. Original data are generated on \( 21 \times 376, 16 \times 352, 17 \times 376, \) and \( 16 \times 576 \) unstructured mesh respectively, and corresponding inverse problem are solved on \( 9600, 7392, 7392, 7392 \) unstructured mesh. Four point sources are placed in \( (-20, 0), (0, 20), (20, 0) \) and \( (0, -20) \). Iterative relative errors are defined by

\[
\epsilon_{\mu_a} = \frac{\|\mu_a - \mu_a^*\|_2}{\|\mu_a^*\|_2}, \quad \epsilon_{\mu_s} = \frac{\|\mu_s - \mu_s^*\|_2}{\|\mu_s^*\|_2}.
\]

4.2.1 Reconstruction of \( \mu_a \) given \( \mu_s \) Using one measurement with boundary source in \( (20, 0) \), we apply the improved fixed-point iterative and BB methods to retrieve absorption coefficients of templates, given scattering coefficients. Our initial guesses are set to be the same as background. Figure 3 shows results after 50 iterations for noiseless data. Then \( 5\% \) Gaussian noise (i.e. \( \epsilon = 5\% \)) is added to data according to (43), and the final reconstructed images are showed in figure 4. Specific iterative relative errors are showed in figure 5. From figures 3 and 4, it seems that both methods can retrieve absorption coefficient with almost equally accurate solutions in the cases of noise-free and noisy measurements. However, from figure 5, it is obvious that improved fixed-point iteration converges more rapidly than the BB method. Improved fixed-point iteration achieves the critical point only after a few iterations.
To attain the same accuracy, the number of iteration of improved fixed-point iteration is about the half that of the BB method. Owing to solving the adjoint RTEs in the BB method, the improved fixed-point iteration can achieve the same accuracy as the BB method in about a quarter the computational time of the latter.

4.2.2. Reconstruction of $\mu_a$ and $\mu_s$ simultaneously. We apply the BB method to reconstruct absorption and scattering coefficients simultaneously from four measurements. The boundary point sources are placed at the top, bottom, left and right sides. The reconstruction results from noiseless data are showed in figure 6. Corresponding reconstruction results from data added by 5% Gaussian noise are shown in figure 7. The relative $L^2$ errors of reconstructions of optical coefficients of four templates are tabulated in table 1. From figures 6 and 7, we can see that BB method retrieves absorption coefficient accurately in piecewise constant and smooth cases, and even is stable to noisy data. In the first, third and
Figure 5. Specific iterative relative errors $\epsilon_{\mu_c}$ of $\mu_c$, where ‘BB’ and ‘IFPI’ mean the results of BB method and improved fixed-point iteration respectively.

Figure 6. Reconstructions of optical coefficients by BB method from noiseless data. Top row: reconstruction of $\mu_a$. Bottom row: reconstruction of $\mu_s$. 
fourth templates, some jagged borders of inclusions are reasonable, which are caused by the interpolation of data from fine mesh to coarse mesh. Nevertheless, the reconstruction of scattering coefficient is not satisfactory. In the piecewise constant case, the edges between different pieces are blurred, which can be seen from the first, third, and fourth template in figures 6 and 7. This is due to the insensitive scattering coefficient. After all, there is no explicit $\mu_s$ in the formula (6).

5. Conclusion

In this paper, we investigate the reconstruction of absorption and scattering coefficients in QPAT in two cases. For a given scattering coefficient, we propose an improved fixed-point iterative method to reconstruct the absorption coefficient and prove its convergence. The advantages of this reconstruction algorithm are its fast convergence, and that it does not require the initial guess to be close to the exact solution. Additionally, it does not need to solve adjoint RTE in the optimization approach. For the simultaneous reconstruction of the two coefficients, we apply a state-of-the-art BB method, which does not need linesearch to compute stepsize. Indeed, linesearch involves solving RTE several times, which is expensive. The BB method only uses the values of the previous two steps to update its current estimation. Numerical results show that improved fixed-point iteration can achieve almost the same accuracy as the BB method. It takes about a quarter of the computational time of the BB method. Moreover,
the algorithm is stable to noise. In the unknown scattering coefficient case, numerical results show that the BB method can obtain quite accurate absorption estimates. However, the result of the estimation of the scattering coefficient is not satisfactory. There is no explicit $\mu_s$ in (6), so it is insensitive to the objective function. There is obvious blur near the borders in piecewise constant templates.

In future, we will design a better error functional, to improve the reconstruction of $\mu_s$. Although the BB method can numerically recover absorption and scattering coefficients simultaneously, its convergence remains to be studied. Further, for more practical application, we will do some realistic simulations, such as 3D templates or real physical QPAT data.

Acknowledgments

We would like to thank the anonymous referees for their useful comments that helped us improve the quality of the paper. We would like to thank Prof Markus Haltmeier (Department of Mathematics, University of Innsbruck Technikestraße) and Dr Hao Gao (Wallace H Coulter Department of Biomedical Engineering, Georgia Institute of Technology) for their helpful advice on RTE solver. We also thank Mr Ji Li (School of Mathematical Sciences, Peking University) for helpful comments on this manuscript draft. This work was supported by NSF grants of China (61421062, 11471024).

Appendix A. Convergence of the solver for original RTE

We use the norm $\| \phi \| := \sqrt{\sum_{k=0}^{P-1} \omega_k \int_{\Omega} \phi_k^2}$ to estimate the accuracy of the numerical solution, where $\omega_k$ is defined by $\omega_k = \int_{S^{n-1}} L_k(\theta) d\theta$ with Lagrangian function $L_k(\theta)$ in $S^{n-1}$.

We denote the solution of original RTE (2) and angular discretized equation (34) by $\phi(x, \theta)$ and $[\phi] := (\phi_k(x))_{k=0}^{P-1}$ respectively, and assume the angular and spatial mesh size are $h_0$ and $h$ respectively. Then there are some convergence results.

**Theorem 7 ([16]).** Assume $\phi(x, \theta) \in C^2(\Omega \times S^{n-1})$, with vacuum boundary condition and sufficiently fine angular mesh

$$\| \phi - [\phi] \| = \sum_{k=0}^{P-1} \omega_k \int_{\Omega} (\phi_k(x) - \phi(x, \theta_k))^2 \, dx = O(h_0^2). \quad (A.1)$$

The convergence of spatial discrete scheme (35) is discussed next. Assume triangulation $T_h$ with $h = \sup_{S \in T_h} \text{diam}(S)$, $V^d_h := \{ v : v|_{S \in T_h} \in P^d(S), v|_{\Gamma_h \setminus S} = 0 \}$, where $P^d(S)$ is the space of $d$-degree polynomials. Let $(\mu, v)_S := \int_S \mu v \, dx$, $\Gamma_h := (\cup_{S \in T_h} \partial S) \setminus \partial \Omega$ and $\partial \Omega_{\mu, v, (+)} := \{ x \in \partial \Omega : \nu \cdot \theta_k \leq (>)0 \}$. Then the convergence result is as follows.

**Theorem 8 ([16]).** With vacuum boundary condition, if $\phi_h^k \in V^d_h$ satisfies

$$A_k(\phi_h^k, v) = (q, v) + (q_h, v)_S + (\phi_h^k, v)_S + (\phi_h^k, v)_S + (\phi_h^k, v)_S + (\phi_h^k, v)_S + (\phi_h^k, v)_S + (\phi_h^k, v)_S + (\phi_h^k, v)_S,$$

where

$$A_k(\phi_h^k, v) = \sum_{s} (\theta_k \cdot \nabla \phi_h^k + (\mu_h + \mu_s)\phi_h^k - \mu_s \sum_{s'} \omega_{s,s'} \phi_h^{s'}, v) + (\phi_h^k, v)_S + (\phi_h^k, v)_S + (\phi_h^k, v)_S + (\phi_h^k, v)_S + (\phi_h^k, v)_S + (\phi_h^k, v)_S + (\phi_h^k, v)_S + (\phi_h^k, v)_S + (\phi_h^k, v)_S,$$
with \( \phi^\pm_h = \lim_{\epsilon \to 0^+} \phi^\pm_h(x + \epsilon \theta_k) \), \( \langle u, v \rangle_{T^*_h} = \int_{T^*_h} u \theta_k \cdot v \) and \( \langle u, v \rangle_{\partial T^*_h} = \int_{\partial T^*_h} u \theta_k \cdot v \), then \( \phi^h := (\phi^h) \in (V_d^h)^p \) satisfies
\[
\| [\hat{\phi}] - \phi^h \| \leq C h^{d+1/2} |\hat{\phi}|_{d+1},
\]  
where the \( |\phi|_{d+1} := \sum_{k=0}^{P-1} \omega_k |\phi_k|^2_{d+1} \).

Combining (A.1) and (A.3), when \( d = 1 \) we can obtain the estimate
\[
[\phi] - \phi^h = O(h^2) + O(h^{3/2}).
\]

Obviously, the corresponding update scheme of (A.2) is (35) when \( d = 1 \). Therefore, the convergence proof of (35) is complete.

**Appendix B. Convergence of the solver for adjoint RTE**

As for adjoint RTE (30), we denote the solution of it and its angular discretized equation
\[
\begin{align*}
-\theta_k \cdot \nabla \tilde{\phi}_k + (\mu_a + \mu_s) \tilde{\phi}_k &= \mu_s \sum_{k'=1}^{P} \omega_{k'} \tilde{\phi}_{k'} + q_k, 1 \leq k \leq P
\end{align*}
\]
by \( \tilde{\phi}(x, \theta) \) and \( [\tilde{\phi}] := (\tilde{\phi}_k(x))_{k'=0}^{P-1} \) respectively. Through similar discussion, we obtain following results.

**Theorem 9.** Assume \( \tilde{\phi}(x, \theta) \in C^2(\Omega \times \mathcal{S}^{n-1}) \), with vacuum boundary condition and sufficiently fine angular mesh
\[
\| [\tilde{\phi}] - \tilde{\phi} \| = \sqrt{\sum_{k=0}^{P-1} \omega_k \int_{\Omega} (\tilde{\phi}_k(x) - \tilde{\phi}(x, \theta_k))^2 \, dx} = O(h^2).
\]

**Theorem 10.** With vacuum boundary condition, if \( \tilde{\phi}_k^h \in V_d^h \) satisfies
\[
\begin{align*}
A_k(\tilde{\phi}_k^h, v) &= q(v) + \langle q_h, v \rangle_{\partial \Omega^T} \quad \forall v \in V_d^h, k = 0, 1, \ldots, P - 1,
\end{align*}
\]
where
\[
A_k(\tilde{\phi}_k^h, v) = \sum_{k} (-\theta_k \cdot \nabla \tilde{\phi}_k^h + (\mu_a + \mu_s) \tilde{\phi}_k^h - \mu_s \sum_{k'} \omega_{k'} \tilde{\phi}_{k'}^h, v) + \langle \tilde{\phi}_k^h, v \rangle_{T^*_h} + \langle \tilde{\phi}_k^h, v \rangle_{\partial \Omega^T},
\]
then \( \tilde{\phi}^h := (\tilde{\phi}^h) \in (V_d^h)^p \) satisfies
\[
\| [\tilde{\phi}] - \tilde{\phi}^h \| \leq C h^{d+1/2} |\tilde{\phi}|_{d+1},
\]
where \( |\tilde{\phi}|_{d+1} := \sum_{k=0}^{P-1} \omega_k |\tilde{\phi}_k|^2_{d+1} \).

Combining (B.2) and (B.4), when \( d = 1 \) we can obtain the estimate
\[
\| \tilde{\phi} - \tilde{\phi}^h \| = O(h^2) + O(h^{3/2}).
\]

Obviously, the corresponding update scheme of (B.3) is (42) when \( d = 1 \). Therefore, the convergence proof of (42) is complete.
Appendix C. Comments on the log-type error functional

According to [27], taking logs accelerates the convergence of the minimization method significantly. This is due to taking logs changing the contour shape of the error functional. From figure 2 in [27], the contours of the error functional become less narrow after taking logs. In numerical simulations, we find the log-type error functional decreasing more quickly. Indeed, we can explain it by a one-dimension example. Let \( f(x) \) be a continuous function on bounded region \([a, b]\), which satisfies \( 0 < f(x) < 1 \). Given \( h^* = f(x^*) \) (\( \forall x \in [a, b] \)), we can apply the least squares method to recover \( x^* \). Define error functional

\[
\mathcal{J}_1(x) = \frac{1}{2} \|f(x) - h^*\|^2_2,
\]

\[
\mathcal{J}_2(x) = \frac{1}{2} \|\log(f(x)) - \log(h^*)\|^2_2.
\]

Then the gradients are

\[
\nabla \mathcal{J}_1(x) = f'(x)(f(x) - h^*),
\]

\[
\nabla \mathcal{J}_2(x) = \frac{f'(x)}{f(x)}(\log(f(x)) - \log(h^*)).
\]

**Theorem 11.** For error functionals defined in (C.1), we claim that there exists \( \delta > 0 \) such that if \( x \in (x^* - \delta, x^* + \delta) \), \( \|\nabla \mathcal{J}_1(x)\| \leq \|\nabla \mathcal{J}_2(x)\| \).

**Proof.** Assume there exists some \( \delta_0 \) such that if \( x \in (x^* - \delta_0, x^* + \delta_0) \), \( f(x) \neq h^* \). Otherwise, \( |\nabla \mathcal{J}_1| = |\nabla \mathcal{J}_2| \) in \( (x^* - \delta_0, x^* + \delta_0) \).

Since \( f(x) \) is continuous, \( f(x) \to f(x^*) \) as \( x \to x^* \), then

\[
\lim_{x \to x^*} \log \left(1 + \frac{f(x) - f(x^*)}{f(x^*)} \right)^{\frac{f(x^*)}{\|\nabla \mathcal{J}_2\|}} = 1.
\]

Therefore, for arbitrary \( \epsilon > 0 \), there exists some \( \delta_1 > 0 \) such that if \( x \in (x^* - \delta_1, x^* + \delta_1) \),

\[
1 - \epsilon < \left| \log \left(1 + \frac{f(x) - f(x^*)}{f(x^*)} \right)^{\frac{f(x^*)}{\|\nabla \mathcal{J}_2\|}} \right| < 1 + \epsilon. \tag{C.3}
\]

Dividing (C.3) by \( f(x^*) \), we can obtain

\[
\left| \frac{1 - \epsilon}{f(x^*)} \right| < \left| \log \left(1 + \frac{f(x) - f(x^*)}{f(x^*)} \right)^{\frac{f(x^*)}{\|\nabla \mathcal{J}_2\|}} \right| < \left| \frac{1 + \epsilon}{f(x^*)} \right|.
\]

Let \( 0 < \epsilon \leq 1 - f(x^*) \), we can get

\[
f(x) < 1 < \frac{1}{|f(x) - f(x^*)|} \left| \log \left(1 + \frac{f(x) - f(x^*)}{f(x^*)} \right) \right|.
\]

Clearly, it follows that

\[
|f'(x)(f(x) - f(x^*))| \leq \left| \frac{f'(x)}{f(x)}(\log(f(x)) - \log(f(x^*))) \right|. \tag{C.4}
\]

Let \( \delta = \min\{\delta_0, \delta_1\} \), then \( |\nabla \mathcal{J}_1| \leq |\nabla \mathcal{J}_2| \) in \( (x^* - \delta, x^* + \delta) \).
Theorem 12. For error functionals defined in (C.1), we claim that there exists $\delta > 0$ such that functional $\mathcal{D}(x) := J_2 - J_1$

is monotonically increasing with respect to $\alpha(x) := |x - x^*|$ in domain $[x^* - \delta, x^* + \delta]$.  

**Proof.** Assume there exists $\delta_0$ such that $f(x) \neq h^*$ in $[x^* - \delta_0, x^* + \delta_0]$. Since $f$ is continuous, there exists some $\delta_1 < \delta_0$ such that $|f(x) - h^*|$ is increasing with respect to $\alpha(x)$, that is, in interval $[x^* - \delta_1, x^* + \delta_1]$, the closer between $x$ and $x^*$ then the smaller $|f(x) - h^*|$. For arbitrary $x = x^* + \Delta x \in [x^* - \delta_1, x^* + \delta_1]$, let $f(x) = h^* + \Delta f$. Then we have

$$
\mathcal{D}(x) = \frac{1}{2} \|\log f(x) - \log h^*\|_2^2 - \frac{1}{2} \|f(x) - h^*\|_2^2 \\
= \frac{1}{2} (\log f + \log h^* - h^*)(\log f - \log h^* + h^*) \\
= \frac{1}{2} (\log(1 + \frac{\Delta f}{h^*}) + \Delta f)(\log(1 + \frac{\Delta f}{h^*}) - \Delta f).
$$

Clearly, $\log(1 + \frac{\Delta f}{h^*}) + \Delta f$ and $\log(1 + \frac{\Delta f}{h^*}) - \Delta f$ are monotonically increasing with respect to $\Delta f$. Combing $\log(1 + \frac{\Delta f}{h^*}) + \Delta f = \log(1 + \frac{\Delta f}{h^*}) - \Delta f = 0$ when $\Delta f = 0$, we obtain $\mathcal{D}(x)$ is non-negative and monotonically increasing with respect to $|\Delta f|$. Therefore, $\mathcal{D}$ is monotonically increasing with respect to $\alpha(x)$.

Let $\delta = \min\{\delta_0, \delta_1\}$; this completes the proof. $$\square$$

From theorems 11 and 12, we see if $\sup f(x) < 1$, taking logs not only makes the error functional steeper near the minimizer for some fixed direction, but also the farther between $x$ and $x^*$, the more significant is the change from $J_1(x)$ to $J_2(x)$. Then, for multivariate functions, taking logs makes the contours less narrow. Thence, the log-type error functional improves the convergence of the minimization method. As for the general case of $\sup f(x) > 1$, we can replace $f(x)$ by $f(x)/c$ for some constant $c > \sup f(x)$.

**References**

[1] Agranovsky M, Kuchment P and Kunyansky L 2009 On reconstruction formulas and algorithms for the thermoacoustic tomography pp 89–101 (Informa UK Limited)
[2] Bal G, Jollivet A and Jugnon V 2010 Inverse transport theory of photoacoustics Inverse Problems 26 025011
[3] Bal G and Ren K 2011 Multi-source quantitative photoacoustic tomography in a diffusive regime Inverse Problems 27 075003
[4] Bal G and Ren K 2012 On multi-spectral quantitative photoacoustic tomography in diffusive regime Inverse Problems 28 025010
[5] Bal G and Uhlmann G 2010 Inverse diffusion theory of photoacoustics Inverse Problems 26 85010
[6] Barzilai J and Borwein J M 1988 Two-point step size gradient methods IMA J. Numer. Anal. 8 141–8
[7] Beretta E, Muszkieta M, Naeter W and Scherzer O 2016 A variational method for quantitative photoacoustic tomography with piecewise constant coefficients (Walter de Gruyter GmbH)
[8] Case K M and Zweifel P F 1963 Existence and uniqueness theorems for the neutron transport equation J. Math. Phys. 4 1376–85
[9] Cox B T, Arridge S R and Beard P C 2009 Estimating chromophore distributions from multiwavelength photoacoustic images J. Opt. Soc. Am. A 26 443–55
[10] Cox B, Lauffer J G, Arridge S R and Beard P C 2012 Quantitative spectroscopic photoacoustic imaging: a review J. Biomed. Opt. 17 061202
[11] Cox B T, Arridge S R, Küstli K P and Beard P C 2006 Two-dimensional quantitative photoacoustic image reconstruction of absorption distributions in scattering media by use of a simple iterative method Appl. Opt. 45 1866
[12] Ding T, Ren K and Vallélian S 2015 A one-step reconstruction algorithm for quantitative photoacoustic imaging Inverse Problems 31 095005
[13] Feng J and Gao H 2016 Limited view multi-source photoacoustic tomography with finite transducer dimension Photons Plus Ultrasound: Imaging and Sensing 2016 vol 9708 (Bellingham, WA: SPIE/SpIE-Intl Soc Optical Eng) p 97080Y
[14] Gao H, Osher S and Zhao H 2012 Quant. Photoacoust. Tomogr. 2035 131–58
[15] Gao H and Zhao H 2009 A fast forward solver of radiative transfer Transp. Theory Stat. Phys. 38 149–92
[16] Gao H and Zhao H 2012 Analysis of a numerical solver for radiative transport equation Math. Comput. 82 153–72
[17] Grippo L and Sciandrone M 2002 Nonmonotone globalization techniques for Barzilai–Borwein gradient method Comput. Optim. Appl. 23 143–69
[18] Haltmeier M, Neumann L and Rabanser S 2015 Single-stage reconstruction algorithm for quantitative photoacoustic tomography Inverse Problems 31 065005
[19] Harrison T, Shao P and Zemp R J 2013 A least-squares fixed-point iterative algorithm for multiple illumination photoacoustic tomography Biomed. Opt. Express 4 2224–30
[20] Kuchment P and Kunyansky L 2011 Mathematics of Photoacoustic and Thermoacoustic Tomography (Berlin: Springer) pp 817–65
[21] Lv T and Zhou T 2014 Variational iterative algorithms in photoacoustic tomography with variable sound speed J. Comput. Math. 32 579–600
[22] Mamonov A V and Ren K 2014 Quantitative photoacoustic imaging in radiative transport regime Common. Math. Sci. 12 201–34
[23] Pulkkinen A, Cox B T, Arridge S R, Kaipio J P and Tarvainen T 2014 A Bayesian approach to spectral quantitative photoacoustic tomography Inverse Problems 30 065012
[24] Pulkkinen A, Cox B T, Arridge S R, Goh H, Kaipio J P and Tarvainen T 2016 Direct estimation of optical parameters from photoacoustic time series in quantitative photoacoustic tomography IEEE Trans. Med. Imaging 35 2497–508
[25] Saratoon T, Tarvainen T, Cox B T and Arridge S R 2013 A gradient-based method for quantitative photoacoustic tomography using the radiative transfer equation Inverse Problems 29 075006
[26] Shao P, Cox B and Zemp R J 2011 Estimating optical absorption, scattering, and Grueneisen distributions with multiple-illumination photoacoustic tomography Appl. Opt. 50 3145–54
[27] Tarvainen T, Cox B T, Kaipio J P and Arridge S R 2012 Reconstructing absorption and scattering distributions in quantitative photoacoustic tomography Inverse Problems 28 084009
[28] Xu M and Wang L V 2002 Time-domain reconstruction for thermoacoustic tomography in a spherical geometry IEEE Trans. Med. Imaging 21 814–22
[29] Xu M and Wang L V 2006 Photoacoustic imaging in biomedicine Rev. Sci. Instrum. 77 041101
[30] Yao L, Sun Y and Jiang H 2010 Transport-based quantitative photoacoustic tomography: simulations and experiments Phys. Med. Biol. 55 1917–34
[31] Zhang X, Zhou W, Zhang X and Gao H 2014 Forward backward splitting method for quantitative photoacoustic tomography Inverse Problems 30 125012