INSTITUTE OF PHYSICS PUBLISHING INVERSE PROBLEMS

Inverse Problems **22** (2006) 1659–1675 doi:10.1088/0266-5611/22/5/008

# Mathematical theory and numerical analysis of bioluminescence tomography

**Weimin Han**[1], **Wenxiang Cong**[2] **and Ge Wang**[3]

[1] Department of Mathematics, University of Iowa, Iowa City, IA 52242, USA
[2] Bioluminescence Tomography Laboratory, Department of Radiology, University of Iowa, Iowa City, IA 52242, USA
[3] Bioluminescence Tomography Laboratory, Department of Radiology and Department of Mathematics, University of Iowa, Iowa City, IA 52242, USA

E-mail: whan@math.uiowa.edu, wenxiang-cong@uiowa.edu and ge-wang@ieee.org

Received 14 January 2006, in final form 23 June 2006
Published 21 August 2006
Online at stacks.iop.org/IP/22/1659

**Abstract**

Molecular imaging is widely recognized as the main stream in the next generation of biomedical imaging. Bioluminescence tomography (BLT) is a rapidly developing new area of molecular imaging. The goal of BLT is to provide quantitative three-dimensional reconstruction of a bioluminescent source distribution within a small animal from optical signals on the surface of the animal body. In this paper, a mathematical framework is established for BLT. Solution existence and uniqueness are established. Continuous dependence of the solution is demonstrated with respect to data. Stable BLT schemes are studied, leading to error estimates and convergence of the methods. A numerical example is presented to illustrate the algorithmic performance.

(Some figures in this article are in colour only in the electronic version)

## 1. Introduction

In the post-genomic era, great efforts are being made to link genes to phenotypic expressions for development of molecular medicine. An important component of this perspective is small animal imaging that allows *in vivo* studies at anatomical, functional, cellular and molecular levels. In molecular/cellular imaging, small animal organs and tissues are often labelled with reporter probes that generate detectable signals [13, 23]. This approach is already widely used to investigate tumorigenesis, cancer metastasis, cardiac diseases, cystic fibrosis, gene therapies and so on. Despite the availability of traditional imaging modalities including computed tomography (CT), positron emission tomography (PET) and magnetic resonance imaging (MRI) [15], optical imaging methods such as florescence molecular tomography (FMT) [16] and bioluminescent imaging (BLI) [17] are becoming increasingly important, because these techniques directly reveal molecular and cellular activities sensitively, specifically and cost effectively [7].

Our Iowa team conceptualized and developed the first bioluminescence tomography (BLT) prototype which compensates for heterogeneous scattering properties of a mouse and performs quantitative 3D reconstruction of internal sources from bioluminescent views measured on the external surface of the mouse [6, 21, 22]. BLT has now become a rapidly developing area for molecular imaging. The introduction of BLT relative to planar bioluminescent imaging (BLI) can be in a substantial sense compared to the development of x-ray CT based on radiography. Without BLT, bioluminescent imaging is primarily qualitative. With BLT, quantitative and localized analyses on a bioluminescent source distribution become feasible inside a living mouse. In the March 2005 issue of the Molecular Imaging Outlook (http://www.diagnosticimaging.com/molecularimagingoutlook/2005mar/02.jhtml), Contag mentioned that BLI arose out of the frustration with sampling limitations of the standard assay techniques. Also, since the genes are duplicated with the cell division, BLI is more sensitive than other techniques such as nuclear imaging in which the radioactive signal will be reduced with the cell division. Piwnica-Worms underlined in the same magazine that BLI could be applied to study almost all diseases in every small animal model.

The pre-requisites for this imaging project are bioluminescent probes, substrate administration and subsequent signal collection. Naturally occurring luciferases exhibit emission maxima between 480 nm and 635 nm. In principle, we may use luciferases with different spectral properties to sense various biological events. Recent developments in luciferase technology have confirmed spectrally shifted signals from luciferases in various species and/or by mutagenesis. Among the current methods, combining firefly (*Photinus pyralis*) ($\lambda_{\max} = 562$ nm) and click beetle (*Pyrophorus plagiopthalamus*) luciferase ($\lambda_{\max} = 615$ nm) is an attractive option because they utilize the same non-toxic substrate. There are also areas for further development of bioluminescence reporters that could expand the utility of bioluminescent imaging. These include isolation of novel luciferases, mutation of known luciferases, luminescence–resonance energy transfer to red-emitting fluorescent proteins and development of luciferase substrate analogues with different emission properties. Coincidentally, the latest development in the cooled-CCD camera technology has reached the point that allows us to detect very weak optical signals such as bioluminescent signals on the body surface of a mouse.

Let $\Omega$ be a domain in $\mathbb{R}^3$ with a Lipschitz boundary $\Gamma$, $q$ a bioluminescent source function in $\Omega$ and $u(\boldsymbol{x}, \boldsymbol{\theta}, t)$ the radiance in $\boldsymbol{\theta} \in S^2$ ($S^2$: the unit sphere) at $\boldsymbol{x} \in \Omega$. The radiative transfer equation (RTE) [15] can be used to describe the bioluminescent photon transport in the medium as follows:

$$\frac{1}{c}\frac{\partial u}{\partial t} + \boldsymbol{\theta}\cdot\nabla_{\boldsymbol{x}} u + \mu u = \mu_s \int_{S^2} \eta(\boldsymbol{\theta}\cdot\boldsymbol{\theta}')u(\boldsymbol{x}, \boldsymbol{\theta}', t)\,\mathrm{d}\theta' + q,$$

where $c$ denotes the photon speed, $\mu = \mu_a + \mu_s$ with $\mu_a$ and $\mu_s$ being the absorption and scattering coefficients, and the scattering kernel $\eta$ satisfies $\int_{S^2}\eta(\boldsymbol{\theta}\cdot\boldsymbol{\theta}')\,\mathrm{d}\theta' = 1$. Mathematically, BLT is the source inversion problem that is to recover $q$ from optical measurement on the domain boundary $\Gamma$, utilizing detailed knowledge on the optical properties of $\Omega$. Note that obtaining the individualized spatially variant optical properties is critical for BLT to work effectively.

Because the RTE is difficult to handle and because in the range of around 600 nm photon scattering outperforms absorption in a mouse, usually a diffusion approximation of the RTE is employed [15]. The steady-state form of the diffusion approximation is the following boundary value problem (BVP):

$$-\mathrm{div}(D\nabla u_0) + \mu_a u_0 = q_0 \qquad \text{in} \quad \Omega, \tag{1.1}$$

$$u_0 + 2D\frac{\partial u_0}{\partial \nu} = g^- \qquad \text{on} \quad \Gamma, \tag{1.2}$$

where $u_0(\boldsymbol{x}) = \int_{S^2} u(\boldsymbol{x}, \boldsymbol{\theta})\, \mathrm{d}\theta$, $g^-$ is the incoming flux on $\Gamma$, $D = 1/[3(\mu_a + \mu_s')]$, $\mu_s' = (1 - \overline{\eta})\mu_s$, $\overline{\eta} = \int_{S^2} \boldsymbol{\theta} \cdot \boldsymbol{\theta}' \eta(\boldsymbol{\theta} \cdot \boldsymbol{\theta}')\, \mathrm{d}\theta'$, $q_0(\boldsymbol{x}) = \int_{S^2} q(\boldsymbol{x}, \boldsymbol{\theta})\, \mathrm{d}\theta/(4\pi)$ and $\partial/\partial\nu$ denotes the outward normal derivative on $\Gamma$. The measurement is

$$g = -D\frac{\partial u}{\partial \nu} \qquad \text{on} \quad \Gamma. \tag{1.3}$$

In this paper, we study the BLT problem of finding a source function $q_0$ given $g^-$ and $g$ such that (1.1), (1.2) and (1.3) are satisfied. We call this the pointwise formulation [22]. Inverse source problems in such a pointwise formulation are the subject of numerous references. A recent reference is [8], where the objective is to identify the source function as a linear combination of monopolar and dipolar sources. Note that the BLT problem is fundamentally different from the so-called diffuse optical tomography (DOT) problem. Using the diffusion approximation, the DOT problem is to find optical properties (absorption and reduced scattering coefficients) of an object from diffuse signals generated by a controllable optical stimulation and measured on the external surface of the object. In other words, in the BLT problem the source is unknown, while in the DOT problem the optical properties are to be determined. Theoretical studies on the solution non-uniqueness of the DOT problem were reported in [1, 2, 11, 19].

This paper provides a mathematical and numerical analysis of the BLT problem. In section 2, we point out the ill-posedness of the pointwise formulation: (1) in general, there are infinite many solutions; (2) when the form of the source function is specified, generally there are no solutions. Moreover, it is easy to see that the source function does not depend continuously on the data. In section 3, we establish a comprehensive mathematical framework for the BLT problem through Tikhonov regularization; we demonstrate the solution existence, uniqueness and continuous dependence on the data. In section 4, we introduce stable numerical methods for the BLT reconstruction, derive error estimates and show convergence of the numerical solutions. In section 5, we include a numerical example to show the performance of the numerical methods.

## 2. Ill-posedness of the pointwise formulation

To avoid complicated subscripts, we simplify the notation by expressing the BLT problem as the determination of a source function $p$ in the differential equation

$$-\mathrm{div}(D\nabla u) + \mu u = p \qquad \text{in} \quad \Omega \tag{2.1}$$

from two boundary conditions:

$$u + 2D\frac{\partial u}{\partial \nu} = g^- \qquad \text{on} \quad \Gamma, \tag{2.2}$$

$$D\frac{\partial u}{\partial \nu} = -g \qquad \text{on} \quad \Gamma. \tag{2.3}$$

Here $D = [3(\mu + \mu')]^{-1}$, $\mu$ and $\mu'$ are given absorption and scattering coefficients, the influx $g^-$ is a given function and is zero in a typical BLT problem, whereas $g$ is the measurement.

From (2.2) and (2.3) we obtain a third possible boundary condition

$$u = g^- + 2g \qquad \text{on} \quad \Gamma. \tag{2.4}$$

Only two of the three boundary conditions (2.2)–(2.4) are independent. To determine the source function $p$, we may associate one of the three boundary conditions (2.2), (2.3) or (2.4) with the differential equation (2.1) to form a boundary value problem, and choose one of the remaining boundary conditions to form the inverse problem for $p$. Thus, there are six possibilities leading to the inverse problem. To be definite, we choose (2.3) as the boundary condition for the boundary value problem, and use (2.4) for the recovery of the source function $p$. In other words, we study the following problem. Discussions of the other five possible inverse problems are similar.

**Problem 2.1.** *Given $D > 0$, $\mu \geqslant 0$, $g_1$ and $g_2$, suitably smooth, find a source function $p$ such that the solution of the boundary value problem*

$$-\mathrm{div}(D\nabla u) + \mu u = p \qquad \text{in} \quad \Omega, \tag{2.5}$$

$$D\frac{\partial u}{\partial \nu} = g_2 \qquad \text{on} \quad \Gamma \tag{2.6}$$

*satisfies*

$$u = g_1 \qquad \text{on} \quad \Gamma. \tag{2.7}$$

Let us point out that problem (2.1) is ill-posed due to its pointwise formulation. To illustrate this, we first show that solution uniqueness does not hold. We recall a trace theorem in the theory of Sobolev spaces [12] (the symbol $\gamma$ stands for the trace operator).

**Theorem 2.2.** *Let $\Omega \subset \mathbb{R}^d$ be an open bounded set with a $C^{1,1}$ boundary $\Gamma$. Then the operator: $v \mapsto Rv = \{\gamma v, \gamma\partial v/\partial\nu\}$, defined for $v \in C^{1,1}(\overline{\Omega})$, has a unique continuous extension as an operator from $H^2(\Omega)$ onto $H^{3/2}(\Gamma) \times H^{1/2}(\Gamma)$. This operator has a right continuous inverse.*

We also have the next result.

**Proposition 2.3.** *Let $\Omega \subset \mathbb{R}^d$ be an open bounded set with a $C^{1,1}$ boundary $\Gamma$ and $u \in H^2(\Omega)$. Then there are infinitely many functions $v \in H^2(\Omega)$ such that*

$$\gamma v = \gamma u, \qquad \gamma\frac{\partial v}{\partial \nu} = \gamma\frac{\partial u}{\partial \nu}. \tag{2.8}$$

**Proof.** Let $\Omega_1$ be a proper subset of $\Omega$, and let $\Omega_2 \neq \emptyset$ be a proper subset of $\Omega_1$. Let $\phi \in C^\infty(\overline{\Omega})$ be a cut-off function such that $\phi(\boldsymbol{x}) = 1$ for $\boldsymbol{x} \in \Omega\backslash\Omega_1$, $\phi(\boldsymbol{x}) = 0$ for $\boldsymbol{x} \in \Omega_2$. Then $v = u\phi \in H^2(\Omega)$ and satisfies the two equalities in (2.8). Since there are infinitely many choices for $\phi$, the statement of the proposition is valid. □

As a consequence of theorem 2.2 and proposition 2.3, the following result holds.

**Corollary 2.4.** *Let $g_1 \in H^{3/2}(\Gamma)$, $g_2 \in H^{1/2}(\Gamma)$. Then problem (2.1) has infinitely many solutions.*

**Proof.** By theorem 2.2, a solution of problem (2.1) exists. By proposition 2.3, there are infinitely many solutions. □

Since there is no solution uniqueness, it is then natural to seek a solution $p$ of a particular form as in [22]. However, this generally leads to non-solvability of the inverse problem. For instance, let $p_0$ be a particular function (e.g. $p_0 = \chi_B$ the characteristic function of a targeted subset $B$ of $\Omega$) and we seek a source function solution of the BLT problem in the form $p = \lambda p_0$ with $\lambda$, a scalar parameter. Let $u_0$ be the solution of the boundary value problem:

$$-\mathrm{div}(D\nabla u_0) + \mu u_0 = p_0 \quad \text{in} \quad \Omega, \qquad u_0 + 2D\frac{\partial u_0}{\partial \nu} = 0 \quad \text{on} \quad \Gamma,$$

where we used the boundary condition of the form (2.2) with zero influx. Then the solution of the boundary value problem

$$-\mathrm{div}(D\nabla u) + \mu u = \lambda p_0 \quad \text{in} \quad \Omega, \qquad u + 2D\frac{\partial u}{\partial \nu} = 0 \quad \text{on} \quad \Gamma$$

is $u = \lambda u_0$. For this $u$ to satisfy the measurement equation (2.3), we need $\lambda u_0 = 2g$ on $\Gamma$. Consequently, a source function solution exists if and only if $g(\boldsymbol{x})/u_0(\boldsymbol{x})$ is a constant for $\boldsymbol{x} \in \Gamma$. Since $g$ and $u_0$ are obtained from different sources, we do not expect $g(\boldsymbol{x})/u_0(\boldsymbol{x})$ to be a constant. In other words, by restricting the form of the source function $p$, the BLT problem generally fails to have a solution.

Even when the solution existence and uniqueness issues could be settled, the pointwise formulation is not appropriate for practical purpose, since the source function solution does not continuously depend on the measurement.

It is then natural to study the BLT problem from a different perspective.

## 3. The inverse problem through regularization

For the given data, we assume $\Omega \subset \mathbb{R}^d$ $(d \leqslant 3)$ is a non-empty, open, bounded set with a Lipschitz boundary $\Gamma$, $D \in L^\infty(\Omega)$, $D \geqslant D_0$ a.e. in $\Omega$ for some constant $D_0 > 0$, $\mu \in L^\infty(\Omega)$, $\mu \geqslant 0$ a.e. in $\Omega$, and $\mu$ is positive over a subset of $\Omega$ with positive measure. Also assume $g_1 \in L^2(\Gamma)$, $g_2 \in L^2(\Gamma)$. For the formulation to be broad enough, so as to cover the practically important situation of compactly supported source functions for example, we consider, instead of (2.5), the following differential equation:

$$-\mathrm{div}(D\nabla u) + \mu u = p\chi_{\Omega_0} \quad \text{in} \quad \Omega.$$

Here $\Omega_0$ is a measurable subset of $\Omega$ ($\Omega_0 = \Omega$ is allowed), $\chi_{\Omega_0}$ is the characteristic function of $\Omega_0$, i.e., its value is 1 in $\Omega_0$, and is 0 in $\Omega\backslash\Omega_0$. Thus, the light source exists only in $\Omega_0$, known as the permissible region. Note that the subset $\Omega_0$ itself can be the union of a collection of disjoint subsets of $\Omega$. Suppose we seek the source function $p$ in a closed convex subset $Q_{ad}$ of the space $L^2(\Omega_0)$. Examples include $Q_{ad} = L^2(\Omega_0)$, or the subset of $L^2(\Omega_0)$ of non-negatively valued functions, or a finite-dimensional subspace or subset of linear combinations of specified functions such as the characteristic functions of certain subsets of $\Omega$. For any $q \in L^2(\Omega_0)$, denote $u = u(q) \in H^1(\Omega)$ the solution of the problem

$$\int_\Omega (D\nabla u \cdot \nabla v + \mu u v)\,\mathrm{d}x = \int_{\Omega_0} q v\,\mathrm{d}x + \int_\Gamma g_2 v\,\mathrm{d}s \qquad \forall\, v \in H^1(\Omega). \quad (3.1)$$

By the well-known Lax–Milgram lemma (e.g. [3, 10]), due to the assumptions made on the data, the solution $u(q)$ exists and is unique.

Following the idea of Tikhonov regularization (e.g. [9, 20]), we let

$$J_\varepsilon(q) = \|u(q) - g_1\|^2_{L^2(\Gamma)} + \varepsilon\|q\|^2_{L^2(\Omega_0)}, \qquad \varepsilon \geqslant 0, \quad (3.2)$$

and introduce the following BLT problem.

**Problem 3.1.** *Find $p_\varepsilon \in Q_{ad}$ such that*

$$J_\varepsilon(p_\varepsilon) = \inf_{q \in Q_{ad}} J_\varepsilon(q). \quad (3.3)$$

We comment that mathematically, it is more natural to use $\|u(q) - g_1\|^2_{H^{1/2}(\Gamma)}$ to replace $\|u(q) - g_1\|^2_{L^2(\Gamma)}$ in definition (3.2). However, for actual simulation, it is more convenient to use the $\|\cdot\|_{L^2(\Gamma)}$ norm in the objective function. Inclusion of the term $\varepsilon\|q\|^2_{L^2(\Omega_0)}$ in (3.2) is to stabilize the solution through maintaining the size in $L^2(\Omega_0)$. This term can be replaced by other stabilizing terms appropriate for concrete applications. When $g_1$ is known only on a portion $\Gamma_0$ of $\Gamma$, we need to replace $\|u(q) - g_1\|^2_{L^2(\Gamma)}$ by $\|u(q) - g_1\|^2_{L^2(\Gamma_0)}$ in (3.2).

In the study of problem (3.1), we will use a few properties of the boundary value problem (3.1) and the objective functional $J_\varepsilon(\cdot)$. For $p, q \in L^2(\Omega_0)$, it is easy to see that $u(p + q) - u(p)$ is linear in $q$, and

$$J'_\varepsilon(p)q = 2(u(p) - g_1, u(p + q) - u(p))_{L^2(\Gamma)} + 2\varepsilon(p, q)_{L^2(\Omega_0)},$$
$$J''_\varepsilon(p)q^2 = 2\|u(p + q) - u(p)\|^2_{L^2(\Gamma)} + 2\varepsilon\|q\|^2_{L^2(\Omega_0)}.$$

Hence, for $\varepsilon > 0$, $J_\varepsilon$ is strictly convex. Also,

$$u(p_1 + q) - u(p_2) = u(p_1) - u(p_2 - q) \qquad \forall\, p_1, p_2, q \in L^2(\Omega_0).$$

We now address the existence and uniqueness issue.

**Theorem 3.2.** *For any $\varepsilon > 0$, problem (3.1) has a unique solution $p_\varepsilon \in Q_{ad}$. Moreover, the solution $p_\varepsilon \in Q_{ad}$ is characterized by a variational inequality*

$$(u(p_\varepsilon) - g_1, u(q) - u(p_\varepsilon))_{L^2(\Gamma)} + \varepsilon(p_\varepsilon, q - p_\varepsilon)_{L^2(\Omega_0)} \geqslant 0 \qquad \forall\, q \in Q_{ad}. \tag{3.4}$$

*When $Q_{ad} \subset L^2(\Omega_0)$ is a subspace, the inequality is reduced to a variation equation*

$$(u(p_\varepsilon) - g_1, u(q) - u(0))_{L^2(\Gamma)} + \varepsilon(p_\varepsilon, q)_{L^2(\Omega_0)} = 0 \qquad \forall\, q \in Q_{ad}. \tag{3.5}$$

**Proof.** Note that $L^2(\Omega_0)$ is a Hilbert space, $Q_{ad} \subset L^2(\Omega_0)$ is convex and closed, $J_\varepsilon : Q_{ad} \to \mathbb{R}$ is strictly convex, continuous and coercive, i.e., $J_\varepsilon(q) \to \infty$ as $\|q\|_{L^2(\Omega_0)} \to \infty$. By a standard result on convex minimization (see, e.g. [3, theorem 3.3.12]), there is a unique solution $p_\varepsilon \in Q_{ad}$ to problem (3.1), and the solution is characterized by the relation (see [3, theorem 5.3.19])

$$J'_\varepsilon(p_\varepsilon)(q - p_\varepsilon) \geqslant 0 \qquad \forall\, q \in Q_{ad},$$

i.e., relation (3.4).

Now assume $Q_{ad} \subset L^2(\Omega_0)$ is a subspace. Take $q = 0$ and $2p_\varepsilon$ in (3.4) to conclude that

$$(u(p_\varepsilon) - g_1, u(0) - u(p_\varepsilon))_{L^2(\Gamma)} + \varepsilon(p_\varepsilon, -p_\varepsilon)_{L^2(\Omega_0)} = 0.$$

So (3.4) is equivalent to

$$(u(p_\varepsilon) - g_1, u(q) - u(0))_{L^2(\Gamma)} + \varepsilon(p_\varepsilon, q)_{L^2(\Omega_0)} \geqslant 0 \qquad \forall\, q \in Q_{ad}. \tag{3.6}$$

Replace $q$ by $-q$ to get

$$(u(p_\varepsilon) - g_1, u(q) - u(0))_{L^2(\Gamma)} + \varepsilon(p_\varepsilon, q)_{L^2(\Omega_0)} \leqslant 0 \qquad \forall\, q \in Q_{ad}. \tag{3.7}$$

Obviously, inequalities (3.6) and (3.7) are equivalent to equality (3.5). □

We then consider the continuous dependence of the solution on the data.

**Theorem 3.3.** *The solution $p_\varepsilon$ of problem (3.1) depends continuously on $g_1 \in L^2(\Gamma)$ and $\varepsilon > 0$.*

**Proof.** Let $\varepsilon > 0$ and $g_1 \in L^2(\Gamma)$ be fixed, and let $\delta > 0$ with $|\delta| \leqslant \varepsilon/2$ and $h_1 \in L^2(\Gamma)$. Denote by $p_{\varepsilon+\delta}$ the solution of problem (3.1) with $\varepsilon$ and $g_1$ replaced by $\varepsilon + \delta$ and $g_1 + h_1$, respectively. Then from (3.4), we have

$$(u(p_{\varepsilon+\delta}) - (g_1 + h_1), u(q) - u(p_{\varepsilon+\delta}))_{L^2(\Gamma)} + (\varepsilon + \delta)(p_{\varepsilon+\delta}, q - p_{\varepsilon+\delta})_{L^2(\Omega_0)} \geqslant 0 \qquad \forall\, q \in Q_{ad}.$$

Choose $q = p_\varepsilon$ in this inequality, choose $q = p_{\varepsilon+\delta}$ in (3.4) and add the two resulting inequalities to obtain

$$\|u(p_{\varepsilon+\delta}) - u(p_\varepsilon)\|^2_{L^2(\Gamma)} + (\varepsilon+\delta)\|p_{\varepsilon+\delta} - p_\varepsilon\|^2_{L^2(\Omega_0)} \leqslant (h_1, u(p_{\varepsilon+\delta}) - u(p_\varepsilon))_{L^2(\Gamma)} - \delta(p_\varepsilon, p_{\varepsilon+\delta} - p_\varepsilon)_{L^2(\Omega_0)}. \tag{3.8}$$

Since $|\delta| \leqslant \varepsilon/2$, we have

$$(\varepsilon+\delta)\|p_{\varepsilon+\delta} - p_\varepsilon\|^2_{L^2(\Omega_0)} \geqslant \frac{\varepsilon}{2}\|p_{\varepsilon+\delta} - p_\varepsilon\|^2_{L^2(\Omega_0)},$$

$$(h_1, u(p_{\varepsilon+\delta}) - u(p_\varepsilon))_{L^2(\Gamma)} \leqslant \frac{1}{2}\|h_1\|^2_{L^2(\Gamma)} + \frac{1}{2}\|u(p_{\varepsilon+\delta}) - u(p_\varepsilon)\|^2_{L^2(\Gamma)},$$

$$-\delta(p_\varepsilon, p_{\varepsilon+\delta} - p_\varepsilon)_{L^2(\Omega_0)} \leqslant \frac{|\delta|}{2}\|p_\varepsilon\|^2_{L^2(\Omega_0)} + \frac{\varepsilon}{4}\|p_{\varepsilon+\delta} - p_\varepsilon\|^2_{L^2(\Omega_0)}.$$

Using these relations in (3.8),

$$\|u(p_{\varepsilon+\delta}) - u(p_\varepsilon)\|^2_{L^2(\Gamma)} + \frac{\varepsilon}{2}\|p_{\varepsilon+\delta} - p_\varepsilon\|^2_{L^2(\Omega_0)} \leqslant \|h_1\|^2_{L^2(\Gamma)} + |\delta|\|p_\varepsilon\|^2_{L^2(\Omega_0)}.$$

From this inequality, we conclude the continuous dependence of the solution on the data. □

Actually, it can be shown that both $p_\varepsilon$ and $u(p_\varepsilon)$ also depend continuously on the coefficient functions $D$ and $\mu$.

We now explore the solution behaviour when $\varepsilon \to \infty$ or $\varepsilon \to 0+$.

**Proposition 3.4.** *If* $0 \in Q_{ad}$, *then* $\|p_\varepsilon\|_{L^2(\Omega_0)} \to 0$ *as* $\varepsilon \to \infty$.

This result follows from the inequality

$$\varepsilon\|p_\varepsilon\|^2_{L^2(\Omega_0)} \leqslant J_\varepsilon(p_\varepsilon) \leqslant J_\varepsilon(0) = \|u(0) - g_1\|^2_{L^2(\Gamma)}.$$

In the case $\varepsilon = 0$, a solution $p \in Q_{ad}$ of problem (3.1) is characterized by the inequality

$$(u(p) - g_1, u(q) - u(p))_{L^2(\Gamma)} \geqslant 0 \qquad \forall q \in Q_{ad}. \tag{3.9}$$

Its proof is similar to that of (3.4). Denote by $S_0 \subset Q_{ad}$ the solution set of problem (3.1). As in [14], the following result holds.

**Proposition 3.5.** *Assume* $S_0$ *is non-empty. Then* $S_0$ *is closed and convex. Moreover,*

$$p_\varepsilon \to p_0 \qquad \text{in} \quad L^2(\Omega_0), \qquad \text{as} \quad \varepsilon \to 0, \tag{3.10}$$

*where* $p_0 \in S_0$ *is the solution of problem (3.1) for* $\varepsilon = 0$ *with minimal* $L^2(\Omega_0)$ *norm:*

$$\|p_0\|_{L^2(\Omega_0)} = \inf_{q \in S_0} \|q\|_{L^2(\Omega_0)}. \tag{3.11}$$

**Proof.** It is straightforward to show that $S_0$ is closed and convex. Here we only prove (3.10). The element $p_0$ of (3.11) exists and is unique (e.g. [3]). We take $q = p_0$ in (3.4), $q = p_\varepsilon$ in (3.9) for $p = p_0$ and adding these two inequalities,

$$\varepsilon(p_\varepsilon, p_0 - p_\varepsilon)_{L^2(\Omega_0)} \geqslant \|u(p_\varepsilon) - u(p_0)\|^2_{L^2(\Gamma)}.$$

Thus, $(p_\varepsilon, p_0 - p_\varepsilon)_{L^2(\Omega_0)} \geqslant 0$, $\|p_\varepsilon\|_{L^2(\Omega_0)} \leqslant \|p_0\|_{L^2(\Omega_0)}$ and $\{p_\varepsilon\}$ is uniformly bounded. Let $\{p_{\varepsilon'}\}$ be a subsequence of $\{p_\varepsilon\}$, converging weakly to $p$. Since $S_0$ is weakly closed, $p \in S_0$. Moreover,

$$\|p\|_{L^2(\Omega_0)} \leqslant \liminf_{\varepsilon' \to 0} \|p_{\varepsilon'}\|_{L^2(\Omega_0)} \leqslant \|p_0\|_{L^2(\Omega_0)}.$$

Since $p_0$ is the unique element in $S_0$ with minimal $L^2(\Omega_0)$ norm, $p = p_0$. Thus the limit $p = p_0$ does not depend on the subsequence selected; consequently, the entire family $p_\varepsilon$ converges weakly to $p_0$ in $L^2(\Omega_0)$ as $\varepsilon \to 0$. Strong convergence is shown as follows:

$$\|p_\varepsilon - p_0\|^2_{L^2(\Omega_0)} = \|p_\varepsilon\|^2_{L^2(\Omega_0)} - 2(p_\varepsilon, p_0)_{L^2(\Omega_0)} + \|p_0\|^2_{L^2(\Omega_0)}$$
$$\leqslant 2\|p_0\|^2_{L^2(\Omega_0)} - 2(p_\varepsilon, p_0)_{L^2(\Omega_0)} \to 0$$

as $\varepsilon \to 0$. □

**Corollary 3.6.** *Suppose the solution set $S_0 = \{p\}$ is a singleton. Then*

$$p_\varepsilon \to p \qquad \text{in} \quad L^2(\Omega), \qquad \text{as} \quad \varepsilon \to 0.$$

Note that $p_0 \in S_0$ is characterized as the unique solution of the variational inequality

$$(p_0, q - p_0)_{L^2(\Omega_0)} \geqslant 0 \qquad \forall\, q \in S_0.$$

We comment that if $Q_{ad}$ is a bounded set, then $S_0$ is non-empty. This follows from applying a standard result on convex minimization, e.g. [3, theorem 3.3.12]. Without further information on $Q_{ad}$, we cannot ascertain uniqueness of a solution when $\varepsilon = 0$.

The solution set $S_0 \neq \emptyset$ also when $\Omega_0 = \Omega \in C^{1,1}$, $g_1 \in H^{1/2}(\Gamma)$ and $Q_{ad} = L^2(\Omega)$, for in this case, we can choose $q \in L^2(\Omega)$ such that the corresponding solution of the boundary value problem (3.1) takes on $g_1$ on the boundary and thus $J_\varepsilon(q)|_{\varepsilon=0} = 0$ is minimally possible.

## 4. Numerical approximations

We now turn to a discussion of numerical solutions of problem (3.1). Let $\{\mathcal{T}_h\}$ ($h$: mesh size) be a regular family of finite-element partitions of $\overline{\Omega}$ such that each element at the boundary $\Gamma$ has at most one non-straight face (for a three-dimensional domain) or side (for a two-dimensional domain). For each triangulation $\mathcal{T}_h = \{K\}$, let $V^h \subset H^1(\Omega)$ be the linear element space. For any $q \in L^2(\Omega_0)$, denote $u^h = u^h(q) \in V^h$ the solution of the problem

$$\int_\Omega (D\nabla u^h \cdot \nabla v^h + \mu u^h v^h)\,\mathrm{d}x = \int_{\Omega_0} q v^h\,\mathrm{d}x + \int_\Gamma g_2 v^h\,\mathrm{d}s \qquad \forall\, v^h \in V^h. \tag{4.1}$$

By the Lax–Milgram lemma, the solution $u^h(q)$ exists and is unique. Let

$$J^h_\varepsilon(q) = \|u^h(q) - g_1\|^2_{L^2(\Gamma)} + \varepsilon\|q\|^2_{L^2(\Omega_0)}, \qquad \varepsilon \geqslant 0. \tag{4.2}$$

The admissible source function space $Q_{ad}$ may or may not need to be discretized. In general, let $Q_{ad,1} \subset Q_{ad}$ be non-empty, closed and convex. Later in the section, we will consider two possible choices of $Q_{ad,1}$. We then introduce the following discretization of problem (3.1).

**Problem 4.1.** *Find $p^h_\varepsilon \in Q_{ad,1}$ such that*

$$J^h_\varepsilon(p^h_\varepsilon) = \inf_{q \in Q_{ad,1}} J^h_\varepsilon(q). \tag{4.3}$$

A discrete analogue of theorem 3.2 and proposition 3.5 is the following result.

**Proposition 4.2.** *For $\varepsilon > 0$, there is a unique solution $p^h_\varepsilon \in Q_{ad,1}$ which is characterized by the discrete variational inequality*

$$\left(u^h(p^h_\varepsilon) - g_1, u^h(q) - u^h(p^h_\varepsilon)\right)_{L^2(\Gamma)} + \varepsilon\left(p^h_\varepsilon, q - p^h_\varepsilon\right)_{L^2(\Omega_0)} \geqslant 0 \qquad \forall\, q \in Q_{ad,1}. \tag{4.4}$$

*The solution $p^h_\varepsilon$ depends continuously on the data. When $Q_{ad,1}$ is a subspace of $L^2(\Omega_0)$, (4.4) reduces to a variational equation:*

$$\left(u^h(p^h_\varepsilon) - g_1, u^h(q) - u^h(0)\right)_{L^2(\Gamma)} + \varepsilon\left(p^h_\varepsilon, q\right)_{L^2(\Omega_0)} = 0 \qquad \forall\, q \in Q_{ad,1}. \tag{4.5}$$

*Suppose problem (4.1) for $\varepsilon = 0$ has a solution, and denote by $S_0^h$ the solution set. Then $S_0^h \subset Q_{ad,1}$ is closed and convex, and $p_\varepsilon^h \to p_0^h$ in $L^2(\Omega_0)$ as $\varepsilon \to 0$, where $p_0^h \in S_0^h$ satisfies*

$$\|p_0^h\|_{L^2(\Omega_0)} = \inf_{q \in S_0^h} \|q\|_{L^2(\Omega_0)}.$$

Similar to problem (3.1), we comment that if $Q_{ad,1}$ is a bounded set, then $S_0^h$ is non-empty. In concrete situations, it is possible to show the non-emptiness of the solution set $S_0^h$ directly.

We now turn to error estimation and convergence analysis of the numerical solutions. For this purpose, we additionally assume

$$\Gamma \in C^{1,1}, \qquad D \in C^1(\overline{\Omega}), \qquad g_2 \in H^{1/2}(\Gamma). \tag{4.6}$$

Then we have the solution regularity bound

$$\|u(q)\|_{H^2(\Omega)} \leqslant c\big(\|q\|_{L^2(\Omega_0)} + \|g_2\|_{H^{1/2}(\Gamma)}\big). \tag{4.7}$$

Such a regularity bound is found in many textbooks on modern PDEs. For example, a proof of this regularity bound is given in [10] under the boundary regularity assumption $\Gamma \in C^2$; a careful examination of the proof there reveals that it is sufficient to assume $\Gamma \in C^{1,1}$. We emphasize that assumptions (4.6) on the data are made to ensure the solution regularity (4.7), which in turn is used in error estimation. Without the solution regularity property, we can still derive error estimates for the numerical solutions, although the error bounds will be of lower orders.

We will make use of the following finite-element interpolation error estimate:

$$\|u - \Pi_{V^h} u\|_{L^2(\Omega)} + h\|u - \Pi_{V^h} u\|_{H^1(\Omega)} \leqslant ch^2\|u\|_{H^2(\Omega)} \qquad \forall\, u \in H^2(\Omega), \tag{4.8}$$

where $\Pi_{V^h} u \in V^h$ is the piecewise linear interpolant of $u$. This error estimate is usually proved when $\Omega$ is a polyhedral/polygonal domain so that each element $K$ in a finite-element partition $\mathcal{T}_h$ has straight faces/sides on its boundary (e.g. [4, 5]). For applications in bioluminescence tomography, $\Omega$ is a smooth domain, and is not polyhedral. In such an application, the error estimate (4.8) still holds. This is argued as follows.

First, recall an extension theorem for functions in Sobolev spaces (e.g. [18, theorem 5, p 181]). The function $u \in H^2(\Omega)$ can be extended to $\tilde{u} \in H^2(\mathbb{R}^d)$ such that for a constant $c$ independent of $\Omega$ and $u$,

$$\|\tilde{u}\|_{H^2(\mathbb{R}^d)} \leqslant c\|u\|_{H^2(\Omega)}. \tag{4.9}$$

Next, for definiteness, consider the case of a two-dimensional domain. Note that the estimate (4.8) is proved by showing its localized version on each element $K$. Consider an element $K$ with a curved side. Denote the two straight sides by $\overline{AB}$ and $\overline{AC}$, and the curved side by $\overset{\frown}{BC}$. We extend $K$ to an element $\tilde{K}$ with straight sides $\overline{AB'}$, $\overline{AC'}$ and $\overline{B'C'}$ in such a way that $B \in \overline{AB'}$, $C \in \overline{AC'}$, $\overline{B'C'} \parallel \overline{BC}$ and $\overline{B'C'} \cap \overset{\frown}{BC} \neq \emptyset$. Then the smallest angle of $\tilde{K}$ is bounded below away from 0, and the diameter of $\tilde{K}$ is $O(h)$. We also view the linear interpolant $\Pi_{V^h} u$ as defined on $\tilde{K}$. The ordinary scaling argument shows that

$$\|\tilde{u} - \Pi_{V^h} u\|_{L^2(\tilde{K})} + h\|\tilde{u} - \Pi_{V^h} u\|_{H^1(\tilde{K})} \leqslant c\|\tilde{u}\|_{H^2(\tilde{K})}.$$

Then we have the error estimate (4.8) with $u$ replaced by $\tilde{u}$. Since $\tilde{u} = u$ in $\Omega$, applying (4.9) we conclude that (4.8) holds.

We start with a preparatory result.

**Lemma 4.3.** *There is a constant $c > 0$ independent of $h$ and $\varepsilon$ such that for any $q \in L^2(\Omega_0)$,*

$$\|u(q) - u^h(q)\|_{L^2(\Gamma)} \leqslant ch^{3/2}\big(\|q\|_{L^2(\Omega_0)} + \|g_2\|_{H^{1/2}(\Gamma)}\big), \tag{4.10}$$

*and for any* $q_1, q_2 \in L^2(\Omega_0)$,

$$\|[u(q_1) - u(q_2)] - [u^h(q_1) - u^h(q_2)]\|_{L^2(\Gamma)} \leqslant ch^{3/2}\|q_1 - q_2\|_{L^2(\Omega_0)}. \quad (4.11)$$

**Proof.** By the Cea's inequality (e.g. [3–5] and (4.8)),

$$\|u(q) - u^h(q)\|_{H^1(\Omega)} \leqslant c \inf_{v^h \in V^h} \|u(q) - v^h\|_{H^1(\Omega)} \leqslant ch\|u(q)\|_{H^2(\Omega)}.$$

Using the solution regularity bound (4.7), we obtain

$$\|u(q) - u^h(q)\|_{H^1(\Omega)} \leqslant ch(\|q\|_{L^2(\Omega_0)} + \|g_2\|_{H^{1/2}(\Gamma)}). \quad (4.12)$$

By applying the Nitsche technique (e.g. [3–5]), we further have

$$\|u(q) - u^h(q)\|_{L^2(\Omega)} \leqslant ch^2\big(\|q\|_{L^2(\Omega_0)} + \|g_2\|_{H^{1/2}(\Gamma)}\big). \quad (4.13)$$

Applying the inequality (deduced from [12, theorem 1.5.1.10])

$$\|v\|_{L^2(\Gamma)} \leqslant c\|v\|^{1/2}_{H^1(\Omega)}\|v\|^{1/2}_{L^2(\Omega)} \qquad \forall\, v \in H^1(\Omega),$$

we obtain (4.10) from (4.12) and (4.13).

Note that $u(q_1) - u(q_2) \in H^1(\Omega)$ is the solution of (3.1) with $q = q_1 - q_2$ and $g_2 = 0$, and $u^h(q_1) - u^h(q_2) \in V^h$ is the corresponding finite-element solution. Then we deduce (4.11) from (4.10). □

To proceed further, we distinguish two cases for $Q_{ad,1}$. In the first case, we take $Q_{ad,1} = Q_{ad}$. This is the natural choice when $Q_{ad}$ is a finite-dimensional subspace or subset of linear combinations of specified functions such as the characteristic functions of certain subsets of $\Omega$. We have the following error bound.

**Theorem 4.4.** *With the choice* $Q_{ad,1} = Q_{ad}$*, there is a constant* $c > 0$ *independent of* $\varepsilon$ *and* $h$ *such that*

$$\big\|u(p_\varepsilon) - u^h\big(p^h_\varepsilon\big)\big\|_{L^2(\Gamma)} + \varepsilon^{1/2}\big\|p_\varepsilon - p^h_\varepsilon\big\|_{L^2(\Omega_0)} \leqslant ch^{3/4}\|u(p_\varepsilon) - g_1\|^{1/2}_{L^2(\Gamma)}\|p_\varepsilon - p^h_\varepsilon\|^{1/2}_{L^2(\Omega_0)}$$
$$+ ch^{3/2}\big(\|p_\varepsilon\|_{L^2(\Omega_0)} + \|g_2\|_{H^{1/2}(\Gamma)}\big). \quad (4.14)$$

**Proof.** We choose $q = p_\varepsilon$ in (4.4),

$$-\big(u^h\big(p^h_\varepsilon\big), u^h(p_\varepsilon) - u^h\big(p^h_\varepsilon\big)\big)_{L^2(\Gamma)} - \varepsilon\big(p^h_\varepsilon, p_\varepsilon - p^h_\varepsilon\big)_{L^2(\Omega_0)} \leqslant -\big(g_1, u^h(p_\varepsilon) - u^h\big(p^h_\varepsilon\big)\big)_{L^2(\Gamma)}.$$

Then,

$$\begin{aligned}
&\big\|u(p_\varepsilon) - u^h\big(p^h_\varepsilon\big)\big\|^2_{L^2(\Gamma)} + \varepsilon\big\|p_\varepsilon - p^h_\varepsilon\big\|^2_{L^2(\Omega_0)}\\
&\quad= \big(u(p_\varepsilon) - u^h\big(p^h_\varepsilon\big), u(p_\varepsilon) - u^h(p_\varepsilon)\big)_{L^2(\Gamma)}\\
&\qquad+ \big(u(p_\varepsilon) - u^h\big(p^h_\varepsilon\big), u^h(p_\varepsilon) - u^h\big(p^h_\varepsilon\big)\big)_{L^2(\Gamma)} + \varepsilon\big(p_\varepsilon - p^h_\varepsilon, p_\varepsilon - p^h_\varepsilon\big)_{L^2(\Omega_0)}\\
&\quad\leqslant \big(u(p_\varepsilon) - u^h\big(p^h_\varepsilon\big), u(p_\varepsilon) - u^h(p_\varepsilon)\big)_{L^2(\Gamma)}\\
&\qquad+ \big(u(p_\varepsilon) - g_1, u^h(p_\varepsilon) - u^h\big(p^h_\varepsilon\big)\big)_{L^2(\Gamma)} + \varepsilon\big(p_\varepsilon, p_\varepsilon - p^h_\varepsilon\big)_{L^2(\Omega_0)}.
\end{aligned}$$

Take $q = p^h_\varepsilon$ in (3.4),

$$0 \leqslant \big(u(p_\varepsilon) - g_1, u\big(p^h_\varepsilon\big) - u(p_\varepsilon)\big)_{L^2(\Gamma)} + \varepsilon\big(p_\varepsilon, p^h_\varepsilon - p_\varepsilon\big)_{L^2(\Omega_0)},$$

and add this inequality to the previous one,

$$\big\|u(p_\varepsilon) - u^h\big(p^h_\varepsilon\big)\big\|^2_{L^2(\Gamma)} + \varepsilon\big\|p_\varepsilon - p^h_\varepsilon\big\|^2_{L^2(\Omega_0)} \leqslant \big(u(p_\varepsilon) - u^h\big(p^h_\varepsilon\big), u(p_\varepsilon) - u^h(p_\varepsilon)\big)_{L^2(\Gamma)}$$
$$+ \big(u(p_\varepsilon) - g_1, u^h(p_\varepsilon) - u(p_\varepsilon) + u\big(p^h_\varepsilon\big) - u^h\big(p^h_\varepsilon\big)\big)_{L^2(\Gamma)}.$$

The first term on the right is bounded by

$$\left\|u(p_\varepsilon)-u^h\left(p_\varepsilon^h\right)\right\|_{L^2(\Gamma)}\|u(p_\varepsilon)-u^h(p_\varepsilon)\|_{L^2(\Gamma)}$$
$$\leqslant \tfrac{1}{2}\left\|u(p_\varepsilon)-u^h\left(p_\varepsilon^h\right)\right\|_{L^2(\Gamma)}^2+\tfrac{1}{2}\|u(p_\varepsilon)-u^h(p_\varepsilon)\|_{L^2(\Gamma)}^2,$$

whereas the second term is bounded by

$$\|u(p_\varepsilon)-g_1\|_{L^2(\Gamma)}\left\|u^h(p_\varepsilon)-u(p_\varepsilon)+u\left(p_\varepsilon^h\right)-u^h\left(p_\varepsilon^h\right)\right\|_{L^2(\Gamma)}$$
$$\leqslant ch^{3/2}\|u(p_\varepsilon)-g_1\|_{L^2(\Gamma)}\left\|p_\varepsilon-p_\varepsilon^h\right\|_{L^2(\Gamma)},$$

where (4.11) was used. Thus,

$$\left\|u(p_\varepsilon)-u^h\left(p_\varepsilon^h\right)\right\|_{L^2(\Gamma)}^2+\varepsilon\left\|p_\varepsilon-p_\varepsilon^h\right\|_{L^2(\Omega_0)}^2$$
$$\leqslant c\|u(p_\varepsilon)-u^h(p_\varepsilon)\|_{L^2(\Gamma)}^2+ch^{3/2}\|u(p_\varepsilon)-g_1\|_{L^2(\Gamma)}\left\|p_\varepsilon-p_\varepsilon^h\right\|_{L^2(\Gamma)}.$$

We then apply (4.10) to deduce (4.14). □

Further error bounds require more information on the data. In this regard, we present two sample results as consequences of theorem 4.4.

**Corollary 4.5.** *Assume $Q_{ad}$ is a bounded set in $L^2(\Omega)$. Then with the choice $Q_{ad,1}=Q_{ad}$, there is a constant $c>0$ independent of $\varepsilon$ and $h$ such that*

$$\left\|u(p_\varepsilon)-u^h\left(p_\varepsilon^h\right)\right\|_{L^2(\Gamma)}+\varepsilon^{1/2}\left\|p_\varepsilon-p_\varepsilon^h\right\|_{L^2(\Omega_0)}\leqslant ch^{3/4}. \tag{4.15}$$

**Proof.** Since $Q_{ad}$ is a bounded set in $L^2(\Omega)$, $\|p_\varepsilon\|_{L^2(\Omega_0)}$ and $\left\|p_\varepsilon^h\right\|_{L^2(\Omega_0)}$ are uniformly bounded with respect to $\varepsilon$ and $h$. By (4.7), $\|u(p_\varepsilon)\|_{H^2(\Omega)}$, and hence $\|u(p_\varepsilon)\|_{L^2(\Gamma)}$ as well, is uniformly bounded. The error bound (4.15) then follows from (4.14). □

We comment that in applications, usually $Q_{ad}$ is a bounded set in $L^\infty(\Omega_0)$. Hence the assumption $Q_{ad}$ being bounded in $L^2(\Omega_0)$ is not restrictive.

For the next sample result, we introduce the following assumption.

Assumption A. The data are such that for some $p_1\in Q_{ad}$, $u(p_1)=g_1$ on $\Gamma$.

As noted at the end of section 3, assumption A is valid when $\Omega_0=\Omega\in C^{1,1}$, $g_1\in H^{1/2}(\Gamma)$ and $Q_{ad}=L^2(\Omega)$. It is also valid when $g_1$ is chosen as the trace of some solution of the boundary value problem (3.1).

**Corollary 4.6.** *Let assumption A hold. Then, with the choice $Q_{ad,1}=Q_{ad}$, there is a constant $c>0$ independent of $\varepsilon$ and $h$ such that*

$$\left\|u(p_\varepsilon)-u^h\left(p_\varepsilon^h\right)\right\|_{L^2(\Gamma)}+\varepsilon^{1/2}\left\|p_\varepsilon-p_\varepsilon^h\right\|_{L^2(\Omega_0)}\leqslant ch^{3/2}. \tag{4.16}$$

**Proof.** From

$$\|u(p_\varepsilon)-g_1\|_{L^2(\Gamma)}^2+\varepsilon\|p_\varepsilon\|_{L^2(\Omega_0)}^2\leqslant J_\varepsilon(p_1)=\varepsilon\|p_1\|_{L^2(\Omega_0)}^2,$$

we see that for all $\varepsilon>0$,

$$\|p_\varepsilon\|_{L^2(\Omega_0)}\leqslant\|p_1\|_{L^2(\Omega_0)},\qquad \|u(p_\varepsilon)-g_1\|_{L^2(\Gamma)}\leqslant\varepsilon^{1/2}\|p_1\|_{L^2(\Omega_0)}. \tag{4.17}$$

Use (4.17) in (4.14) to obtain

$$\left\|u(p_\varepsilon)-u^h\left(p_\varepsilon^h\right)\right\|_{L^2(\Gamma)}+\varepsilon^{1/2}\left\|p_\varepsilon-p_\varepsilon^h\right\|_{L^2(\Omega_0)}\leqslant ch^{3/4}\varepsilon^{1/4}\left\|p_\varepsilon-p_\varepsilon^h\right\|_{L^2(\Omega_0)}^{1/2}+ch^{3/2}. \tag{4.18}$$

The first term on the right-hand side of (4.18) is bounded as follows:

$$ch^{3/4}\varepsilon^{1/4}\left\|p_\varepsilon-p_\varepsilon^h\right\|_{L^2(\Omega_0)}^{1/2}\leqslant\frac{1}{2}\varepsilon^{1/2}\left\|p_\varepsilon-p_\varepsilon^h\right\|_{L^2(\Omega_0)}+\frac{c^2}{2}h^{3/2}.$$

Then from (4.18) we obtain (4.16) (following the convention, $c$ stands for a generic constant, whose value may vary from one place to another). □

Next, we consider another choice of $Q_{ad,1}$. This choice is natural when $Q_{ad}$ is a general subset of $L^2(\Omega_0)$ and so its discretization is necessary. In addition to the regular family of finite-element partitions $\{\mathcal{T}_h\}$ of $\overline{\Omega}$, let $\{\mathcal{T}_{0,H}\}$ be a regular family of finite-element partitions of $\overline{\Omega_0}$ such that each element at the boundary $\partial\Omega_0$ has at most one non-straight face (for a three-dimensional domain) or side (for a two-dimensional domain). The partitions $\mathcal{T}_h$ and $\mathcal{T}_{0,H}$ do not need to be related; however, $\mathcal{T}_h$ is allowed to be constructed on the basis of $\mathcal{T}_{0,H}$. Let $Q^H \subset L^2(\Omega_0)$ be the piecewise constant space. Define $Q_{ad,1} = Q^H_{ad} \equiv Q^H \cap Q_{ad}$. We denote the solution of problem (4.1) by $p^{h,H}_{\varepsilon}$.

Denote by $\Pi^H$ the orthogonal projection operator from $L^2(\Omega_0)$ onto $Q^H$: for $q \in L^2(\Omega_0)$,

$$\Pi^H q \in Q^H, \qquad (\Pi^H q, q^H)_{L^2(\Omega_0)} = (q, q^H)_{L^2(\Omega_0)}, \qquad \forall q^H \in Q^H. \tag{4.19}$$

We have the boundedness inequality

$$\|\Pi^H q\|_{L^2(\Omega_0)} \leqslant \|q\|_{L^2(\Omega_0)}, \qquad \forall q \in L^2(\Omega_0) \tag{4.20}$$

and the error bound (through an argument similar to the one for (4.8))

$$\|q - \Pi^H q\|_{L^2(\Omega_0)} \leqslant cH|q|_{H^1(\Omega_0)} \qquad \forall q \in H^1(\Omega_0). \tag{4.21}$$

It is easy to verify the element-wise formula

$$\Pi^H q|_K = \frac{1}{|K|}\int_K q \,\mathrm{d}x \qquad \forall K \in \mathcal{T}_{0,H}.$$

Consequently, since $Q_{ad} \subset L^2(\Omega_0)$ is convex, $\Pi^H : Q_{ad} \to Q^H_{ad}$, i.e., for $q \in Q_{ad}$, its piecewise constant orthogonal projection $\Pi^H q \in Q^H_{ad}$. We need this property in deriving error estimates below.

**Lemma 4.7.** *There is a constant $c > 0$ independent of $h$ and $H$ such that $\forall q \in L^2(\Omega_0)$,*

$$\|u^h(q) - u^h(\Pi^H q)\|_{H^1(\Omega)} \leqslant cH\|q - \Pi^H q\|_{L^2(\Omega_0)}, \tag{4.22}$$

$$\|u(q) - u^h(\Pi^H q)\|_{H^1(\Omega)} \leqslant cH\|q - \Pi^H q\|_{L^2(\Omega_0)} + ch\big(\|q\|_{L^2(\Omega_0)} + \|g_2\|_{H^{1/2}(\Gamma)}\big). \tag{4.23}$$

**Proof.** Here we only prove (4.22); (4.23) can be proved similarly. Denote $e^{h,H}(q) = u^h(q) - u^h(\Pi^H q) \in V^h$. Then by definition (4.1), we have

$$\int_\Omega [D\nabla e^{h,H}(q)\cdot\nabla v^h + \mu e^{h,H}(q)v^h]\,\mathrm{d}x = \int_{\Omega_0} (q - \Pi^H q)v^h \,\mathrm{d}x, \qquad \forall v^h \in V^h. \tag{4.24}$$

For any $v \in H^1(\Omega)$,

$$\int_{\Omega_0} (q - \Pi^H q)v\,\mathrm{d}x = \int_{\Omega_0} (q - \Pi^H q)(v - r^H)\,\mathrm{d}x, \qquad \forall r^H \in Q^H.$$

Then

$$\int_{\Omega_0} (q - \Pi^H q)v\,\mathrm{d}x \leqslant \|q - \Pi^H q\|_{L^2(\Omega_0)} \inf_{r^H \in Q^H} \|v - r^H\|_{L^2(\Omega_0)}$$
$$\leqslant cH\|q - \Pi^H q\|_{L^2(\Omega_0)}\|v\|_{H^1(\Omega)}.$$

Taking $v^h = e^{h,H}(q)$ in (4.24), we obtain

$$\|e^{h,H}(q)\|_{H^1(\Omega)} \leqslant cH\|q - \Pi^H q\|_{L^2(\Omega_0)},$$

which is (4.22). □

Denote

$$E^H(p_\varepsilon) \equiv \|p_\varepsilon - \Pi^H p_\varepsilon\|_{L^2(\Omega_0)} = \inf_{q^H \in Q^H_{ad}} \|p_\varepsilon - q^H\|_{L^2(\Omega_0)}. \tag{4.25}$$

Note that $E^H(p_\varepsilon)$ is generally expected to be small. We now prove the following error estimate.

**Theorem 4.8.** *With the choice $Q_{ad,1} = Q^H_{ad}$ as a subset of piecewise constant functions, there is a constant $c > 0$ independent of $\varepsilon$, $h$ and $H$ such that*

$$\begin{aligned}\big\|u(p_\varepsilon) - u^h\big(p^{h,H}_\varepsilon\big)\big\|_{L^2(\Gamma)} + \varepsilon^{1/2}\big\|p_\varepsilon - p^{h,H}_\varepsilon\big\|_{L^2(\Omega_0)} &\\ \leqslant c\|u(p_\varepsilon) - g_1\|^{1/2}_{L^2(\Gamma)}\big(H^{1/2}E^H(p_\varepsilon)^{1/2} + h^{3/4}\big\|p_\varepsilon - p^{h,H}_\varepsilon\big\|^{1/2}_{L^2(\Omega_0)}\big)&\\ + cHE^H(p_\varepsilon) + ch\big(\|p_\varepsilon\|_{L^2(\Omega_0)} + \|g_2\|_{H^{1/2}(\Gamma)}\big).&\end{aligned} \tag{4.26}$$

**Proof.** From (4.4) with $q = \Pi^H p_\varepsilon$, we have

$$0 \leqslant \big(u^h\big(p^{h,H}_\varepsilon\big) - g_1, u^h(\Pi^H p_\varepsilon) - u^h\big(p^{h,H}_\varepsilon\big)\big)_{L^2(\Gamma)} + \varepsilon\big(p^{h,H}_\varepsilon, \Pi^H p_\varepsilon - p^{h,H}_\varepsilon\big)_{L^2(\Omega_0)}.$$

Take $q = p^{h,H}_\varepsilon$ in (3.4),

$$0 \leqslant \big(u(p_\varepsilon) - g_1, u\big(p^{h,H}_\varepsilon\big) - u(p_\varepsilon)\big)_{L^2(\Gamma)} + \varepsilon\big(p_\varepsilon, p^{h,H}_\varepsilon - p_\varepsilon\big)_{L^2(\Omega_0)}.$$

Add these two inequalities to obtain

$$\begin{aligned}\big\|u(p_\varepsilon) - u^h\big(p^{h,H}_\varepsilon\big)\big\|^2_{L^2(\Gamma)} + \varepsilon\big\|p_\varepsilon - p^{h,H}_\varepsilon\big\|^2_{L^2(\Omega_0)} &\\ \leqslant \big(u(p_\varepsilon) - u^h\big(p^{h,H}_\varepsilon\big), u(p_\varepsilon) - u^h(\Pi^H p_\varepsilon)\big)_{L^2(\Gamma)}&\\ + \big(u(p_\varepsilon) - g_1, u^h(\Pi^H p_\varepsilon) - u(p_\varepsilon) + u\big(p^{h,H}_\varepsilon\big) - u^h\big(p^{h,H}_\varepsilon\big)\big)_{L^2(\Gamma)}.&\end{aligned} \tag{4.27}$$

In deriving this inequality, we used $\big(p^{h,H}_\varepsilon, \Pi^H p_\varepsilon - p_\varepsilon\big)_{L^2(\Omega_0)} = 0$ following (4.19). The first term on the right-hand side of (4.27) is handled as in the proof of theorem 4.4, followed by an application of (4.23), whereas the second term is bounded by

$$\begin{aligned}\|u(p_\varepsilon) - g_1\|_{L^2(\Gamma)}\big(\|u^h(p_\varepsilon) - u^h(\Pi^H p_\varepsilon)\|_{L^2(\Gamma)}&\\ + \big\|u^h(p_\varepsilon) - u(p_\varepsilon) + u\big(p^{h,H}_\varepsilon\big) - u^h\big(p^{h,H}_\varepsilon\big)\big\|_{L^2(\Gamma)}\big),&\end{aligned}$$

followed by an application of (4.22) and (4.11). The result is (4.26). □

Similar to corollaries 4.5 and 4.6 of theorem 4.4, we have the next two sample results as consequences of theorem 4.8.

**Corollary 4.9.** *Assume $Q_{ad}$ is bounded in $L^2(\Omega_0)$. With the choice $Q_{ad,1} = Q^H_{ad}$ as a subset of piecewise constant functions, there is a constant $c > 0$ independent of $\varepsilon$, $h$ and $H$ such that*

$$\big\|u(p_\varepsilon) - u^h\big(p^{h,H}_\varepsilon\big)\big\|_{L^2(\Gamma)} + \varepsilon^{1/2}\big\|p_\varepsilon - p^{h,H}_\varepsilon\big\|_{L^2(\Omega_0)} \leqslant c(H^{1/2}E^H(p_\varepsilon)^{1/2} + h^{3/4}).$$

**Corollary 4.10.** *Let assumption A hold. With the choice $Q_{ad,1} = Q^H_{ad}$ as a subset of piecewise constant functions, there is a constant $c > 0$ independent of $\varepsilon$, $h$ and $H$ such that*

$$\big\|u(p_\varepsilon) - u^h\big(p^{h,H}_\varepsilon\big)\big\|_{L^2(\Gamma)} + \varepsilon^{1/2}\big\|p_\varepsilon - p^{h,H}_\varepsilon\big\|_{L^2(\Omega_0)} \leqslant c(h + H^{1/2}\varepsilon^{1/4}E^H(p_\varepsilon)^{1/2} + HE^H(p_\varepsilon)).$$

Error bounds of both corollaries 4.9 and 4.10 involve the best approximation error term $E^H(p_\varepsilon)$ defined in (4.25). A crude bound for this term is

$$E^H(p_\varepsilon) \leqslant \|p_\varepsilon\|_{L^2(\Omega_0)} + \|\Pi^H p_\varepsilon\|_{L^2(\Omega_0)} \leqslant 2\|p_\varepsilon\|_{L^2(\Omega_0)} \leqslant c.$$

However, usually we can expect this term to be small as the following result shows.

**Proposition 4.11.** *If $S_0 \neq \emptyset$, then we have the convergence*

$$\|p_\varepsilon - \Pi^H p_\varepsilon\|_{L^2(\Omega_0)} \to 0 \qquad \text{as} \quad H, \varepsilon \to 0.$$

*If $p_\varepsilon \in H^1(\Omega_0)$, then*

$$\|p_\varepsilon - \Pi^H p_\varepsilon\|_{L^2(\Omega_0)} \leqslant cH\|p_\varepsilon\|_{H^1(\Omega_0)}.$$

**Proof.** Since $\|p_\varepsilon - p_0\|_{L^2(\Omega_0)} \to 0$ as $\varepsilon \to 0$ and recalling (4.20), we have

$$\begin{aligned}\|p_\varepsilon - \Pi^H p_\varepsilon\|_{L^2(\Omega_0)} &\leqslant \|p_\varepsilon - p_0\|_{L^2(\Omega_0)} + \|\Pi^H(p_\varepsilon - p_0)\|_{L^2(\Omega_0)} + \|p_0 - \Pi^H p_0\|_{L^2(\Omega_0)} \\ &\leqslant 2\|p_\varepsilon - p_0\|_{L^2(\Omega_0)} + \|p_0 - \Pi^H p_0\|_{L^2(\Omega_0)} \\ &\to 0\end{aligned}$$

as $H, \varepsilon \to 0$. The error bound follows from (4.21). □

We underline that the above theoretical results on the numerical solutions with the second choice of $Q_{ad,1}$ are still valid if $Q^H \subset L^2(\Omega_0)$ is a general finite-element space containing piecewise constants. The proofs of the results are the same as long as we define $\Pi^H$ to be the orthogonal projection operator in $L^2(\Omega_0)$ onto the space of piecewise constants. In actual implementation of the methods, $\overline{\Omega}$ and $\overline{\Omega_0}$ are replaced by their natural polyhedral approximations defined by the finite-element partitions into polyhedral elements. Error bounds for the numerical solutions in such a situation can still be derived, extending the usual arguments as in [5, sections 4.3, 4.4]. Since such a derivation is rather lengthy, we omit it in this paper.

When the regularization parameter $\varepsilon$ is chosen related to the discretization parameters $h$ and $H$, we may express the error bounds in terms of the discretization parameters only. For example, in the context of corollary 4.6, we have

$$\left\|u(p_\varepsilon) - u^h\left(p_\varepsilon^h\right)\right\|_{L^2(\Gamma)} \leqslant ch^{3/2},$$

and if $\varepsilon = ch^\beta, 0 < \beta < 3$, then

$$\left\|p_\varepsilon - p_\varepsilon^h\right\|_{L^2(\Omega_0)} \leqslant ch^{(3-\beta)/2}.$$

Finally, we comment that when the solution set $S_0$ is non-empty, convergence of the numerical solution $p_\varepsilon^h$ to the minimal energy solution $p_0 \in S_0$ follows from the triangle inequality

$$\left\|p_\varepsilon^h - p_0\right\|_{L^2(\Omega_0)} \leqslant \|p_\varepsilon - p_0\|_{L^2(\Omega_0)} + \left\|p_\varepsilon - p_\varepsilon^h\right\|_{L^2(\Omega_0)}$$

together with (3.10) and the convergence of $p_\varepsilon^h$ to $p_\varepsilon$ in $L^2(\Omega_0)$. A similar statement holds for the convergence of $p_\varepsilon^{h,H}$ to $p_0 \in S_0$.

## 5. Numerical example

We report here some numerical results to show the performance of the method. In the example below, we use (2.1) and (2.2) to form the boundary value problem, and reconstruct the source function using the Dirichlet boundary condition (2.4). Since the Robin boundary condition (2.2) is used for defining the boundary value problem, we need to replace (3.1) by

$$\int_\Omega (D\nabla u \cdot \nabla v + \mu u v)\,\mathrm{d}x + \frac{1}{2}\int_\Gamma u v\,\mathrm{d}s = \int_{\Omega_0} q v\,\mathrm{d}x + \frac{1}{2}\int_\Gamma g^- v\,\mathrm{d}s \qquad \forall v \in H^1(\Omega),$$

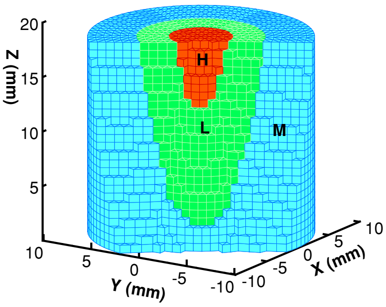


**Figure 1.** The heterogeneous phantom.

and replace (4.1) by

$$\int_\Omega (D\nabla u^h \cdot \nabla v^h + \mu u^h v^h)\,\mathrm{d}x + \frac{1}{2}\int_\Gamma u^h v^h\,\mathrm{d}s = \int_{\Omega_0} q v^h\,\mathrm{d}x + \frac{1}{2}\int_\Gamma g^- v^h\,\mathrm{d}s \qquad \forall v^h \in V^h.$$

All the theoretical results presented in sections 3 and 4 are valid. In the numerical experiment, we let $g^- = 0$.

We do a simulation on a heterogeneous highly scattering phantom. The problem domain is a cylindrical phantom with radius 10 mm and height 26 mm. We set up the coordinate system so that the domain is expressed as

$$\overline{\Omega} = \{\boldsymbol{x} = (x_1, x_2, x_3)^T \in \mathbb{R}^3 \mid x_1^2 + x_2^2 \leqslant 100, 0 \leqslant x_3 \leqslant 26\}.$$

The phantom consists of three kinds of materials, represented by region $M$ (muscle), $L$ (lung) and $H$ (heart), respectively:

$$M = \{\boldsymbol{x} \in \Omega \mid 6 \leqslant \sqrt{x_1^2 + x_2^2} \leqslant 10\},$$
$$L = \{\boldsymbol{x} \in \Omega \mid 3 \leqslant \sqrt{x_1^2 + x_2^2} \leqslant 6\},$$
$$H = \{\boldsymbol{x} \in \Omega \mid \sqrt{x_1^2 + x_2^2} \leqslant 3\}.$$

The optical parameters are assigned to each of the three components as follows:

$$\mu = \begin{cases} 0.020 & \text{in } H, \\ 0.040 & \text{in } L, \\ 0.015 & \text{in } M; \end{cases}$$

$$\mu' = \begin{cases} 1.0 & \text{in } H, \\ 1.5 & \text{in } L, \\ 0.9 & \text{in } M. \end{cases}$$

A spherical light source $p = 3/(4\pi)$ of power 1.0 pW is embedded in the phantom, centred at $(3.975, -1.423, 9.643)^T$ with radius 1 mm. For numerical simulation, we use uniform wedge element partitions of $\overline{\Omega}$, as shown in figure 1. For each partition, denote by $h$ the maximal length of the element edges. The permissible region is chosen to be

$$\Omega_0 = \{\boldsymbol{x} \in \Omega \mid x_1 > 0, 6 < x_3 < 12\}$$

and correspondingly, the admissible set $Q_{ad} = \{q \in L^2(\Omega_0) \mid q \geqslant 0 \text{ a.e. in } \Omega_0\}$. The

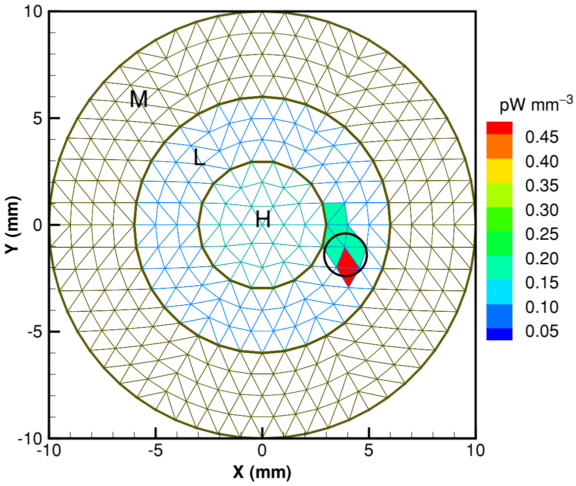


**Figure 2.** Light source reconstruction; $h = 1$.

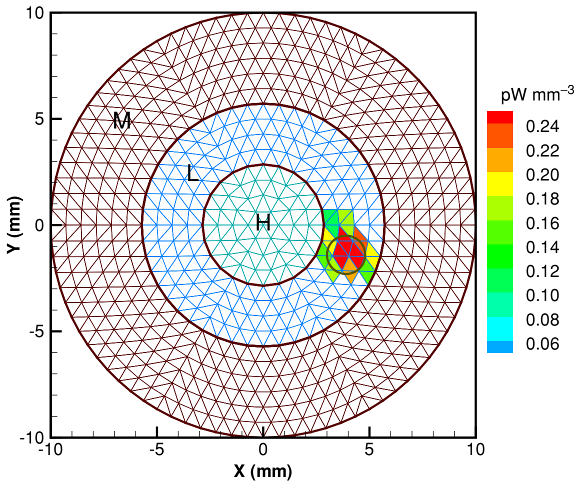


**Figure 3.** Light source reconstruction; $h = 0.7143$.

finite-element partitions are restricted to the permissible region $\Omega_0$ (implying $H = h$) in defining piecewise constant approximations $p_\varepsilon^{h,h}$. We show the reconstructed source function (cross-sectional view at $x_3 = 9.643$) for $h = 1.000$ mm and 0.7143 mm in figures 2 and 3, both with the regularization parameter $\varepsilon = 10^{-7}$. We note the quality improvement in the reconstructed source function as $h$ decreases.

## Acknowledgments

This work was supported by an NIH grant EB001685. We thank the referees for their valuable comments.

## References

[1] Arridg S R 1999 Optical tomography in medical imaging *Inverse Probl.* **15** R41–93
[2] Arridge S R and Lionheart W R B 1998 Nonuniqueness in diffusion-based optical tomography *Opt. Lett.* **23** 882–4
[3] Atkinson K and Han W 2005 *Theoretical Numerical Analysis: A Functional Analysis Framework* (*Texts in Applied Mathematics* vol 39) 2nd edn (New York: Springer)
[4] Brenner S C and Scott L R 2002 *The Mathematical Theory of Finite Element Methods* 2nd edn (New York: Springer)
[5] Ciarlet P G 1978 *The Finite Element Method for Elliptic Problems* (Amsterdam: North-Holland)
[6] Cong W X *et al* 2005 A practical reconstruction method for bioluminescence tomography *Opt. Express* **13** 6756–71
[7] Contag C H and Ross B D 2002 It's not just about anatomy: *in vivo* bioluminescence imaging as an eyepiece into biology *J. Magn. Reson. Imaging* **16** 378–87
[8] El Badia A 2005 Inverse source problem in an anisotropic medium by boundary measurements *Inverse Probl.* **21** 1487–506
[9] Engl H W, Hanke M and Neubauer A 1996 *Regularization of Inverse Problems* (Dordrecht: Kluwer)
[10] Evans L C 1998 *Partial Differential Equations* (Providence, RI: American Mathematical Society)
[11] Gibson A P, Hebden J C and Arridge S R 2005 Recent advances in diffuse optical imaging *Phys. Med. Biol.* **50** R1–R43
[12] Grisvard P 1985 *Elliptic Problems in Nonsmooth Domains* (Boston, MA: Pitman)
[13] Herschman H R 2003 Molecular imaging: looking at problems, seeing solutions *Science* **302** 605–8
[14] Lions J L 1971 *Optimal Control of Systems Governed by Partial Differential Equations* (Berlin: Springer)
[15] Natterer F and Wübbeling F 2001 *Mathematical Methods in Image Reconstruction* (Philadelphia, PA: SIAM)
[16] Niziachristos V *et al* 2002 Fluorescence molecular tomography resolves protease activity *in vivo Nat. Med.* **8** 757–61
[17] Rice B W *et al* 2001 *In vivo* imaging of light-emitting probes *J. Biomed. Opt.* **6** 432–40
[18] Stein E M 1970 *Singular Integrals and Differentiability Properties of Functions* (Princeton, NJ: Princeton University Press)
[19] Taufiquar A T and Yoon J-R 2006 On uniqueness in refractive index optical tomography *Inverse Probl.* **22** L1–5
[20] Tikhonov A N 1963 Regularization of incorrectly posed problems *Sov. Dokl.* **4** 1624–7
[21] Wang G *et al* 2003 Development of the first bioluminescent CT scanner *Radiology* **229(P)** 566
[22] Wang G, Li Y and Jiang M 2004 Uniqueness theorems in bioluminescence tomography *Med. Phys.* **31** 2289–99
[23] Weissleder R and Ntziachristos V 2003 Shedding light onto live molecular targets *Nat. Med.* **9** 123–8