A new class of accelerated regularization methods, with application to bioluminescence tomography

Rongfang Gong\textsuperscript{1}, B. Hofmann\textsuperscript{2} and Ye Zhang\textsuperscript{2,3,‡}

\textsuperscript{1}Department of Mathematics, Nanjing University of Aeronautics and Astronautics, 211106 Nanjing, China

\textsuperscript{2}Faculty of Mathematics, Chemnitz University of Technology, 09107 Chemnitz, Germany

\textsuperscript{3}School of Science and Technology, Örebro University, 70182 Örebro, Sweden

E-mail: grf\_math@nuaa.edu.cn, hofmannb@mathematik.tu-chemnitz.de, ye.zhang@oru.se

Abstract. In this paper we propose a new class of iterative regularization methods for solving ill-posed linear operator equations. The prototype of these iterative regularization methods is in the form of second order evolution equation with a linear vanishing damping term, which can be viewed not only as an extension of the asymptotical regularization, but also as a continuous analog of the Nesterov’s acceleration scheme. New iterative regularization methods are derived from this continuous model in combination with damped symplectic numerical schemes. The regularization property as well as convergence rates and acceleration effects under the Hölder-type source conditions of both continuous and discretized methods are proven.

The second part of this paper is concerned with the application of the newly developed accelerated iterative regularization methods with a posteriori stopping rule to the diffusion-based bioluminescence tomography, which is modeled as an inverse source problem in elliptic partial differential equations with both Dirichlet and Neumann boundary data. Several numerical examples, as well as a comparison with the state-of-the-art methods, are given to show the accuracy and the acceleration effect of the new methods.

1. Introduction

In the first part of this paper we consider linear operator equations

\[ Kf = y, \] 

where \( K \) is a compact linear operator acting between two infinite dimensional Hilbert spaces \( Q \) and \( Y \) such that the range \( \mathcal{R}(K) \) of \( K \) is an infinite dimensional subspace of \( Y \). Then the range \( \mathcal{R}(K) \) is a non-closed subset of \( Y \). For simplicity, we denote by \( \langle \cdot, \cdot \rangle \) and \( \| \cdot \| \) in the sequel the inner products and norms for both Hilbert spaces \( Q \) and \( Y \). The non-closedness of the forward operator \( K \) is typical for operator equations (1) which are models for linear inverse problems. More precisely, due to the compactness of \( K \), the operator equation (1) is ill-posed of type II in the sense of Nashed (cf. [16]). As a consequence of this ill-posedness, a regularization method must be employed in order to obtain reasonable and stable approximate solutions to (1) if the measurement data contains noise. In this context, we consider iterative

‡ Corresponding Author
**Accelerated regularization methods** and assume to know noisy data \( y^\delta \in Y \) instead of the exact right-hand side \( y \in \mathcal{R}(K) \), obeying the deterministic noise model \( \| y^\delta - y \| \leq \delta \) with a priori known noise level \( \delta > 0 \). The focus of our paper will be on studying a specific application in bioluminescence tomography, and we refer to Section 4 for details.

The dominant iterative regularization method for solving (1) should be the Landweber method, given by

\[
f_{k+1}^\delta = f_k^\delta + \Delta t K^*(y^\delta - K f_k^\delta), \quad \Delta t \in (0, 2/\|K^*K\|) \quad (k = 0, 1, 2,...)
\]  

(2)

with some starting element \( f_0 \in Q \), where \( K^* \) denotes the adjoint operator of \( K \). The continuous analog to [2] as \( \Delta t \) tends to zero is known as asymptotic regularization or Showalter’s method (see, e.g., [24, 25]). It is in the form of a first order evolution equation

\[
\dot{f}^\delta(t) + K^*K f^\delta(t) = K^* y^\delta, \quad f(0) = f_0,
\]

(3)

where an artificial scalar time \( t \) is introduced. There must be chosen an appropriate finite stopping time \( T_\ast = T_\ast(\delta) \) (a priori choice) or \( T_\ast = T_\ast(\delta, y^\delta) \) (a posteriori choice) in order to ensure the regularizing property \( f^\delta(T_\ast) \to f^\dagger \) as \( \delta \to 0 \). Here and later on, \( f^\dagger \) represents the unique minimum-norm solution of (1). Moreover, it has been shown that by using Runge-Kutta integrators, all of the properties of asymptotic regularization [3] carry over to its numerical realization [21]. Hence, the continuous model (3) is of particular importance for studying the intrinsic properties of a broad class of general regularization methods for inverse problems, and can be used for the development of new iterative regularization algorithms by combining some appropriate numerical schemes. Inspired by this, the authors in [29] studied the second order asymptotical regularization with the fixed damping parameter.

However, a fatal defect for large-scale problems is the slow performance of the Landweber iteration (too many iterations required for optimal stopping) as well as of the (conventional and second order with a fixed damping parameter) asymptotical regularization methods, i.e. overly excessive stopping times \( T_\ast \) are required for obtaining optimal convergence rates. Therefore, in practice, accelerating strategies are usually used. In so doing, the most commonly known methods are the \( \nu \)-method [6, § 6.3] and the Nesterov acceleration scheme [18]. Recently, the authors in [28] introduced the fractional order asymptotical regularization, and proved its acceleration property. In this paper, we are interested in the following second order evolution equation with a linear vanishing damping term

\[
\ddot{f}^\delta(t) + \frac{1 + 2s}{t} \dot{f}^\delta(t) + K^*K f^\delta(t) = K^* y^\delta, \quad f(0) = f_0, \dot{f}(0) = 0,
\]

(4)

where \( s \geq 0 \) is a fixed number. One motivation to study (4) is that it can be viewed as an infinite dimensional extension of the Nesterov’s scheme in the sense that for all fixed \( T > 0 \) ([23]): \( \lim_{\omega \to 0} \max_{0 \leq k \leq T/\sqrt{s}} \| f_k^\delta - f^\delta(k \sqrt{s}) \|_Q = 0 \), where \( f^\delta(\cdot) \) is the dynamical solution of (4) with \( s \geq 1 \), and \( \{ f_k^\delta \}_k \) is the sequence, generated by the Nesterov’s scheme with parameters \( (\alpha, \omega) \), see formula (135) for details.

It should be noted that the second order dynamic (4) has recently been investigated in [3], where they have proven that the flow (1) with the vanishing initial data yields an optimal regularization method for the linear operator equation (1). In this paper, cf. Section 2, we focus on the acceleration effect in the sense of regularization theory of (1) with an arbitrary initial guess \( f_0 \). The main result of this paper regarding the discretized version of (1) is presented in Section 3, where we demonstrate that by using damped symplectic
integrators, the regularization property and acceleration effect under optimal convergence rates of [1] carry over to its numerical realization. In Section 4, the developed accelerated iterative regularization methods, equipped with a posteriori stopping rule, are applied to a diffusion-based bioluminescence tomography, which can be formulated as:

**Problem 1.** Given $g_1$ and $g_2$, find a bioluminescent source $f$ such that the solution $u$ of the boundary-value problem

\[
\begin{aligned}
-\text{div}(D\nabla u) + \mu u &= f \quad \text{in } \Omega, \\
D\partial_\nu u &= g_2 \quad \text{on } \Gamma
\end{aligned}
\]

satisfies

\[ u = g_1 \quad \text{on } \Gamma. \tag{6} \]

Some numerical examples, as well as a comparison with three well-known existing iterative regularization methods, are presented in Section 5. Concluding remarks are given in Section 6. Some proof details as well as some details on finite element discretization are postponed to the appendices.

### 2. Analysis of the continuous regularization method

#### 2.1. Convergence analysis

We start with the convergence analysis of the continuous method [1] in the sense of regularization theory. Let \( \{\lambda_j; u_j, v_j\}_{j=1}^\infty \) be the well-defined singular system for the compact linear operator $K$, i.e. we have $Ku_j = \lambda_j v_j$ and $Kv_j = \lambda_j u_j$ with ordered singular values $\|K\| = \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_j \geq \lambda_{j+1} \geq \ldots \to 0$ as $j \to \infty$. Since the eigenelements \( \{u_j\}_{j=1}^\infty \) and \( \{v_j\}_{j=1}^\infty \) form complete orthonormal systems (with the exception of null-spaces) in $Q$, the equation in [1] is equivalent to

\[
\langle f^\delta(t), u_j \rangle + \frac{1+2s}{t} \langle f^\delta(t), u_j \rangle + \lambda_j^2 \langle f^\delta(t), u_j \rangle = \lambda_j \langle y^\delta, v_j \rangle, \quad j = 1, 2, \ldots . \tag{7}
\]

Using the decomposition $f^\delta(t) = \sum_j \xi_j(t)u_j$ under the basis \( \{u_j\}_{j=1}^\infty \) in $Q$, we obtain

\[
\ddot{\xi}_j(t) + \frac{1+2s}{t} \dot{\xi}_j(t) + \lambda_j^2 \xi_j = \lambda_j \langle y^\delta, v_j \rangle, \quad j = 1, 2, \ldots . \tag{8}
\]

**Proposition 1.** Let $s \geq -1/2$ be a fixed number. Then, the differential equation [8] with the initial condition $(\dot{\xi}_j(0), \xi_j(0)) = ((f_0, u_j), 0)$ has a unique solution

\[
\xi_j(t) = 2^s \Gamma(s+1) \frac{J_s(\lambda_j t)}{(\lambda_j t)^s} \langle f_0, u_j \rangle + \left(1 - 2^s \Gamma(s+1) \frac{J_s(\lambda_j t)}{(\lambda_j t)^s}\right) \lambda_j^{-1} \langle y^\delta, v_j \rangle,
\]

where $\Gamma(\cdot)$ and $J_s(\cdot)$ denote the gamma function and the Bessel function of first kind of order $s$ respectively.

The proof of the above proposition can be found in Appendix A. By Proposition 1 and the decomposition $f^\delta(t) = \sum_j \xi_j(t)u_j$ we obtain the explicit formula for the solution of [7] as

\[
f^\delta(t) = \sum_j 2^s \Gamma(s+1) \frac{J_s(\lambda_j t)}{(\lambda_j t)^s} \langle f_0, u_j \rangle + \sum_j \left(1 - 2^s \Gamma(s+1) \frac{J_s(\lambda_j t)}{(\lambda_j t)^s}\right) \lambda_j^{-1} \langle y^\delta, v_j \rangle,
\]

\[= (1 - K^* K g(t, K^* K))f_0 + g(t, K^* K)K^* y^\delta, \tag{9}\]

where

\[g(t, \lambda) = \frac{1 - 2^s \Gamma(s+1) \frac{J_s(\sqrt{\lambda t})}{(\sqrt{\lambda t})^s}}{\lambda}. \tag{10}\]
Theorem 1. Let \( f^\delta(t) \) be the dynamic solution of (14). Then, if the terminating time \( T_* = T_*(\delta, y^\delta) \) is chosen so that
\[
\lim_{\delta \to 0} T_* = \infty \quad \text{and} \quad \lim_{\delta \to 0} \delta T_* = 0,
\]
the approximate solution \( f^\delta(T_*) \) converges to \( f^\dagger \) as \( \delta \to 0 \).

Proof. Let \( f(t) \) be the solution of (14) with noise-free data, i.e., \( f(t) = (1 - K^*Kg(t, K^*K))f_0 + g(t, K^*K)K^*y \). Furthermore, define the bias function by
\[
r(t, \lambda) = 1 - \lambda g(t, \lambda) = 2^* \Gamma(s + 1) \frac{J_s(\sqrt{\lambda}t)}{(\sqrt{\lambda}t)^s}.
\]
Obviously, \( r(t, \lambda) \) is the unique solution to
\[
\begin{cases}
\dot{r}(t, \lambda) + \frac{1 + 2\varepsilon}{t} r(t, \lambda) + \lambda r(t, \lambda) = 0, \\
r(0, \lambda) = 1, \quad \dot{r}(0, \lambda) = 0.
\end{cases}
\]
Then, with the help of the intermediate quantity \( f(t) \) and bias function \( r(t, \lambda) \), we obtain the well-known error estimates
\[
\| f^\delta(t) - f^\dagger \| \leq \| f^\delta(t) - f(t) \| + \| f(t) - f^\dagger \|
\leq \delta \sup_{\lambda > 0} \sqrt{\lambda} g(t, \lambda) + \| r(t, K^*K)(f_0 - f^\dagger) \|
\]
by noting that \( y = Kf^\dagger \). Hence, to prove the convergence of the full regularization error, we have to show the convergence of both two terms in the right-hand side of (14).

Let’s first consider the estimate for \( \| r(t, K^*K)(f_0 - f^\dagger) \| \). To this end, define the Lyapunov function of (13) by \( E(t) := \dot{r}^2(t, \lambda) + \lambda r^2(t, \lambda) \). Since
\[
\dot{E}(t) = 2\sqrt{\lambda} \dot{r}(t, \lambda) \dot{r}(t, \lambda) + \lambda r(t, \lambda) = -\frac{2(1 + 2s)}{t} \sqrt{\lambda} r^2(t, \lambda) \leq 0,
\]
\( E(t) \) is a non-increasing function, and consequently, we have \( \lambda r^2(t, \lambda) \leq E(t) \leq E(0) = \lambda \), which implies
\[
|r(t, \lambda)| \leq 1 \text{ for all } \lambda > 0, t \geq 0.
\]
On the other hand, by asymptotic [1] (9.2.1)
\[
J_r(\lambda_j t) = O \left( \sqrt{\frac{2}{\pi \lambda_j t}} \cos \left( \lambda_j t - \frac{\pi}{2} r - \frac{\pi}{4} \right) \right) \text{ as } t \to \infty,
\]
we obtain together with (15) that
\[
\| r(t, K^*K)(f_0 - f^\dagger) \| \to 0 \text{ as } t \to \infty.
\]
Now, consider the quality \( \sqrt{\lambda} g(t, \lambda) \). By the initial condition in (13) and the asymptotic [7] there exists a number \( t_0 \) such that \( r(t, \lambda) \geq 1 - \sqrt{\lambda} t \) for all \( t \in (0, t_0] \). Setting \( t_1 := \min \left\{ 1/2, \sqrt{\lambda t_0}/2 \right\} \), we obtain
\[
r(t, \lambda) \geq 1 - \sqrt{\lambda} t \geq 1 - \frac{\sqrt{\lambda}}{2t_1} t, \text{ for } t \in (0, t_0].
\]
On the other hand, by (15), we deduce that
\[
r(t, \lambda) \geq -1 \geq 1 - \frac{2}{t_0} t \geq 1 - \frac{\sqrt{\lambda}}{2t_1} t \text{ for } t \geq t_0.
\]
Combine (18) and (19) to obtain
\[ \sqrt{\lambda} g(t, \lambda) = \frac{1 - r(t, \lambda)}{\sqrt{\lambda}} \leq \frac{t}{2t_1}, \]  
which implies that \( \delta \sup_{\lambda > 0} \sqrt{\lambda} g(t, \lambda) \leq \delta \frac{1}{2t_1} \to 0 \) under the choice of terminating time in (11).

2.2. Convergence rate and acceleration

The purpose of this subsection is to show that (4) with an appropriate terminating time yields an accelerated optimal regularization method. It is well-known that in order to prove the convergence rate for the approximate solution \( f^\delta(t) \), additional smoothness assumptions on \( f^\dagger \) in correspondence with the forward operator \( K \) have to be fulfilled. For simplicity, we only consider the Hölder type source condition in this paper. By using the technique of the comparison of two qualifications, cf. [15, Def. 2] and [15, Prop. 3, Remark 5 and Lemma 2], the results in this subsection can be easily generalized to general range-type source conditions such as the logarithmic source condition.

Assumption 1. There exists an element \( v_0 \) and numbers \( \mu \in (0, \frac{1 + 2\alpha}{4}] \) and \( \rho \geq 0 \) such that
\[ f_0 - f^\dagger = (K^* K)^\mu v_0 \quad \text{with} \quad \|v_0\| \leq \rho. \]  

Theorem 2. (A priori choice of the terminating time) Under Assumption 1, if the terminating time of the dynamic solution \( f^\delta(T_*) \) of (4) is chosen by \( T_* = \delta^{-\frac{\mu + 1}{2\mu}} \), we have
\[ \|f^\delta(T_*) - f^\dagger\| = O(\delta^{\frac{2\mu}{2\mu + 1}}) \quad \text{as} \quad \delta \to 0. \]  

Proof. According to (16), there exists a pair of numbers \( (C_0, T_0) \) such that for all \( t \geq T_0 \):
\[ J_r(\sqrt{\lambda} t) = C_0 \lambda^{-1/4} t^{-1/2}, \]  
which implies together with (12), (15), and Assumption 1 that
\[ \|r(t, K^* K)(f_0 - f^\dagger)\| = \|r(t, K^* K) (K^* K)^\mu v_0\| \leq \rho \sup_{\lambda \in [0, \|K\|^2]} r(t, \lambda) \lambda^\mu \leq C_1 t^{-2\mu}, \]  
where \( C_1 = C_0 2^{s} (s + 1) \rho \|K\|^{2\mu - \frac{1 + 2\alpha}{4}}. \) We complete the proof by the following inequality
\[ \|f^\delta(T_*) - f^\dagger\| \leq \delta \frac{T_1}{2t_1} + C_1 (T_*)^{-2\mu} = \left(\frac{1}{2t_1} + C_1\right) \delta^{\frac{2\mu}{2\mu + 1}}, \]  
by using (10), (20), (23) and the choice of \( T_* \).}

Now, let us turn to the a posteriori choice of the terminating time \( T_* = T_*(\delta, y^\delta) \). We consider Morozov’s discrepancy principle as the most prominent version exploiting zeros of the discrepancy function
\[ \chi(t) := \|K f^\delta(t) - y^\delta\| - \tau \delta, \]  
where \( \tau > 1 \) is a fixed parameter.

Lemma 1. If \( \|K f_0 - y^\delta\| > \tau \delta \), then \( \chi(T) \) has at least one solution.
The proof of the above lemma can be found in Appendix B. If the function $\chi(T)$ has more than one root, we recommend selecting $T_*$ from the rule

$$\chi(T_*) = 0 < \chi(T), \quad \forall T < T_*.$$  

In other words, $T_*$ is the first time point for which the size of the residual $\|Kf^{\delta}(t) - y^{\delta}\|$ has about the order of the data error. By Lemma 1 such $T_*$ always exists. Furthermore, by the proof of Lemma 1, it is easy to show that $\chi(T)$ is bounded by a decreasing function $\mathcal{E}(t) - \tau\delta = \frac{1}{2}\|\hat{f}^{\delta}(t)\|^2 + \|Kf^{\delta}(t) - y^{\delta}\|^2 - \tau\delta$. Roughly speaking, the trend of $\chi(T)$ is to be a decreasing function, where oscillations may occur.

**Theorem 3. (A posteriori choice of the terminating time)** Under Assumption 1, if the terminating time of the dynamic solution $f^{\delta}(T_*)$ of (4) is chosen as a root of the discrepancy function $\mathcal{B}$, we have for any $\mu \in [0, \frac{1+2\mu}{4}]$ the error estimates

$$T_* = O\left(\delta^{-\frac{1}{2\mu+\tau}}\right), \quad \|f^{\delta}(T_*) - f^1\| = O\left(\delta^{\frac{2\mu}{2\mu+1}}\right) \text{ as } \delta \to 0. \quad (26)$$

**Proof.** By Assumption 1 and the interpolation inequality $\|B^p u\| \leq \|B^q u\|^{p/q}\|u\|^{1-p/q}$, we deduce that

$$\|r(T_*, K^* K)(f_0 - f^1)\| = \|(K^* K)^{\mu+1/2} r(T_*, K^* K)v_0\|$$

$$\leq \|(K^* K)^{\mu+1/2} r(T_*, K^* K)v_0\|^{2\mu/(2\mu+1)} \cdot \|r(T_*, K^* K)v_0\|^{1/(2\mu+1)}$$

$$\leq \|Kr(T_*, K^* K)(f_0 - f^1)\|^{2\mu/(2\mu+1)} \cdot \|r(T_*, K^* K)v_0\|^{1/(2\mu+1)} \quad (27)$$

Since the terminating time $T_*$ is chosen according to the discrepancy principle (25), we derive that

$$\tau\delta = \|Kf^{\delta}(T_*) - y^{\delta}\| = \|Kr(T_*, K^* K)(f_0 - f^1) - r(T_*, K^* K)(y^{\delta} - y)\|$$

$$\geq \|Kr(T_*, K^* K)(f_0 - f^1)\| - \|r(T_*, K^* K)(y^{\delta} - y)\| \quad (28)$$

Now we combine the estimates (27) and (28) to obtain, with the source conditions, that

$$\|r(T_*, K^* K)(f_0 - f^1)\|$$

$$\leq \|Kr(T_*, K^* K)(f_0 - f^1)\|^{2\mu/(2\mu+1)} \cdot \|r(T_*, K^* K)v_0\|^{1/(2\mu+1)}$$

$$\leq (\tau\delta + \|r(T_*, K^* K)(y^{\delta} - y)\|)^{2\mu/(2\mu+1)} \rho^{1/(2\mu+1)}$$

$$\leq (\tau + \delta)^{2\mu/(2\mu+1)} \rho^{1/(2\mu+1)} \delta^{2\mu/(2\mu+1)} \quad (29)$$

On the other hand, in a similar fashion to (28), it is easy to show that

$$\tau\delta \leq \|Kr(T_*, K^* K)(f_0 - f^1)\| + \|r(T_*, K^* K)(y^{\delta} - y)\|$$

$$\leq \|Kr(T_*, K^* K)(f_0 - f^1)\| + \delta.$$  

If we combine the above inequality with Assumption 1 and the inequality (23), we obtain

$$(\tau - 1)\delta \leq \|Kr(T_*, K^* K)(f_0 - f^1)\|$$

$$\leq \|((K^* K)^{\mu+1/2} r(T_*, K^* K)v_0\| \leq C_1 T_*^{-(2\mu+1)},$$

which yields the first estimate in (26). The second estimate in (26) holds according to the inequality (24) and the estimate for $T_*$. This completes the proof. \(\Box\)

**Remark 1.** By Theorems 2 and 3 for the proposed continuous regularization method (4) the optimal convergence rates can be obtained with approximately the square root of iterations than would be needed for the ordinary Landweber iteration (or conventional asymptotical regularization method) [6, § 6.2], which means that our method is an accelerated regularization method. However, similar to most of accelerated regularization methods (e.g. $\nu$-method and
Nesterov’s method, the proposed method 4 also shows a saturation phenomenon; i.e., the optimal convergence rate \( \| f^\delta(T_\ast) - f^\dag \| = \mathcal{O}(\delta^{2/(2s+1)}) \) and the asymptotic \( T_\ast = \mathcal{O}(\delta^{-1/(2s+1)}) \) holds only for \( \mu \in (0, \frac{1+2s}{4}) \).

3. A new class of accelerated iterative regularization methods

The evolution equation (4) with an appropriate numerical discretization scheme for the artificial time variable yields a concrete iterative method. This has motivated us to develop some novel iterative regularization methods based on the continuous method (4). The goal of this section is to realize this idea. To this end, let us start with the simplest discretization scheme – the Euler method, which is defined by (denote as \( q = \dot{f} \))

\[
\begin{align*}
  f^{k+1} &= f^k + \Delta t_k v^k, \\
  q^{k+1} &= q^k + \Delta t_k \left( K^\ast(y^\delta - Af^{k+1}) - \frac{1+2s}{t_{k+1}} q^k \right), \\
  f^0 &= f_0, q^0 = 0.
\end{align*}
\]

By elementary calculations, scheme (30) expresses the form of following three-term semi-iterative method

\[
f^{k+1} = f^k + a_k \left( f^k - f^{k-1} \right) + \omega_k K^\ast(y^\delta - Af^k)
\]

with \( a_k = 1 - \Delta t_k \frac{1+2s}{t_k} \) and \( \omega_k = \Delta t_k \). It is well known that the semi-iterative method (31) with special choice of parameters \((a_k, \omega_k)\), equipped with an appropriate stopping rule, yields an order optimal regularization method with (asymptotically) significantly fewer iterations than the classical Landweber iteration [6, § 6.2].

On the other hand, just as with the Runge-Kutta integrators [21] or the exponential integrators [14] for numerically solving first order equations, the damped symplectic integrators are extremely attractive for solving (4), since the schemes are closely related to the canonical transformations [11], and the trajectories of the discretized second flow are usually more stable for its long-term performance. In this work, we consider the Störmer-Verlet scheme, which takes the form

\[
\begin{align*}
  q^{k+\frac{1}{2}} &= q^k - \frac{\Delta t}{2} \frac{1+2s}{t_k} q^{k+\frac{1}{2}} + \frac{\Delta t}{2} K^\ast(y^\delta - Af^k), \\
  f^{k+1} &= f^k + \Delta t q^{k+\frac{1}{2}}, \\
  q^{k+1} &= q^{k+\frac{1}{2}} - \frac{\Delta t}{2} \frac{1+2s}{t_{k+1}} q^{k+\frac{1}{2}} + \frac{\Delta t}{2} K^\ast(y^\delta - Af^{k+1}), \\
  f^0 &= f_0, q^0 = 0.
\end{align*}
\]

Remark 2. Denote by (SE1) the standard symplectic Euler method of second order system (4). Let (SE2) be the adjoint scheme of (SE1). Then, the scheme (32) follows from the composition (SE2) \( \circ \) (SE1).

Surprisingly, the scheme (32) shares the same recurrence form (31), but with the different parameters

\[
a_k = \frac{1 - \frac{\Delta t(1+2s)}{2t_k}}{1 + \frac{\Delta t(1+2s)}{2t_k}} = \frac{2k - (1 + 2s)}{2k + (1 + 2s)}, \quad \omega_k = \frac{\Delta t^2}{1 + \frac{\Delta t(1+2s)}{2t_k}} = \frac{2\Delta t^2 k}{2k + (1 + 2s)}.
\]

The goal in this section is to prove that the scheme (32) with an appropriate iteration stopping rule yields an accelerated iterative regularization method.
Without loss of generality, define that $\omega_k := \Delta t^2/2$ for $k \leq s + 1/2$. Consequently, $\omega_k \geq \Delta t^2/2$ for all $k \in \mathbb{N}$.

As the Landweber iterates, according to (31), the iterates $f^k$ of (32) obviously belong to the Krylov subspace $\text{Span}\left\{K^*y^\delta, (K^*K)^{k-1}K^*y^\delta\right\}$. Therefore, the solution $f^k$ of (32) can be written as $f^k = g_k(K^*K)^{-1}K^*y^\delta$, where $g_k$ is a polynomial of degree $k - 1$, and the residual polynomials

$$r_k(\lambda) = 1 - \lambda g_k(\lambda)$$

(33)

exhibit the following property.

**Lemma 2.** Let $\Delta t \in (0, \|K\|)$ be a fixed number. Then, the residual polynomials of scheme (32) satisfy the following inequality

$$\sup_{\lambda \in (0, \|K\|^2]} \lambda^\mu r_k(\lambda) \leq c_1 k^{-2\mu}, \quad \text{if } \mu \in (0, 1/2],$$

(34)

and

$$\sup_{\lambda \in (0, \|K\|^2]} \lambda^\mu r_k(\lambda) \leq c_2 k^{-(\mu+1/2)}, \quad \text{if } \mu > 1/2.$$  

(35)

**Proof.** This proof uses the technique in [18]. By elementary calculations, the residual polynomials of (32) satisfy the recurrence relation

$$r_k = (1 - \omega_k \lambda) r_{k-1} + a_k (r_{k-1} - r_{k-2}),$$

which can be rewritten as

$$r_{k+1} = (1 - \omega_{k+1} \lambda) \left[ (1 - \theta_k) r_k + \theta_k \left( r_{k-1} + \frac{1}{\theta_{k-1}} (r_k - r_{k-1}) \right) \right],$$

(36)

where

$$a_k = \frac{\theta_k}{\theta_{k-1}} (1 - \theta_{k-1}).$$

(37)

Let us first show that there exists a sequence $(\theta_k)_k$ such that it satisfies (37) and the following inequalities simultaneously

$$\frac{(1 - \theta_k)^2}{\theta_{k-1}^2} \leq \frac{1}{\theta_{k-1}^2} \quad \text{and} \quad \theta_k \in (0, 1) \text{ for } k > 1.$$  

(38)

According to (37), we have (obviously, $\theta_k \neq 1$ for $k > 1$)

$$\theta_k = \frac{a_k \theta_{k-1}}{1 - \theta_{k-1}}.$$  

(39)

Putting (39) in (38), we obtain

$$\frac{1 - (a_k + 1) \theta_{k-1}}{a_k \theta_{k-1}} \leq \frac{1}{\theta_{k-1}},$$

which implies that one can choose together with the definition of $a_k$ that

$$\theta_k = \frac{1 - a_{k+1}}{1 + a_{k+1}} = \frac{2s + 1}{2k + 2} = O\left(\frac{1}{k}\right).$$  

(40)
Denote by \( \hat{r}_k := r_{k-1} + \frac{1}{\theta_{k-1}} (r_k - r_{k-1}) \). Then, we derive together with (36) and (38) that
\[
\omega_k \lambda r_{k+1}^2 + (1 - \omega_k \lambda) \hat{r}_{k+1}^2 = (1 - \omega_k \lambda) \left[ (1 - \theta_k \frac{r_k}{\theta_{k-1}}) \omega_k \lambda r_k^2 + (1 - \omega_k \lambda) \hat{r}_k^2 \right]
\]
\[
\leq (1 - \omega_k \lambda) \left[ \omega_k \lambda r_k^2 + (1 - \omega_k \lambda) \hat{r}_k^2 \right]
\]
\[
\leq \prod_{i=1}^{k} (1 - \omega_i \lambda) \cdot \left[ \frac{\omega_k}{\theta_{k-1}^2} \lambda r_k^2 + (1 - \omega_k \lambda) \hat{r}_k^2 \right] = \prod_{i=1}^{k+1} (1 - \omega_i \lambda)
\]
by noting that \( \theta_0 = \hat{r}_0 = 1 \) (as \( r_{-1} = r_0 \equiv 1 \)). Inequality (41) immediately yields
\[
\lambda r_{k+1}^2 \leq \frac{\theta_k^2}{\omega_k} \frac{\theta_{k-1}^2}{\omega_k} \prod_{i=1}^{k+1} (1 - \omega_i \lambda) \leq \frac{\theta_k^2}{\Delta t^2} \prod_{i=1}^{k+1} (1 - \Delta t^2 \lambda)^{k+1},
\]
as well as \( |\hat{r}_{k+1}| \leq 1 \). The latter inequality together with the recurrence (36) and initial data \( r_{-1} = r_0 \equiv 1 \) implies
\[
|r_k| \leq 1.
\]
If \( \mu \in (0, 1/2] \), (42) and (43) immediately gives
\[
\lambda^\mu r_k(\lambda) \leq \left( \lambda r_k^2(\lambda) \right)^\mu \leq \frac{2\theta_k^2}{\Delta t^2} \leq c_1 k^{-2\mu}.
\]
If \( \mu > 1/2 \), we obtain together with (42) that
\[
\lambda^\mu r_k(\lambda) = \lambda^\mu r_k(\lambda) \lambda^{-\frac{1}{2}} \leq \frac{\sqrt{2\theta_k}}{\Delta t} \left( 1 - \Delta t^2 \lambda \right)^{\frac{k+1}{2}} \lambda^{-\frac{1}{2}}
\]
\[
\leq \lambda_{\max} = \frac{\sqrt{2\theta_k}}{\Delta t^2 (2\mu + k)} \left( k+1 \right)^{\frac{k+1}{2}} \left( 1 - \Delta t^2 \lambda \right)^{\frac{k+1}{2}} \leq c_2 k^{-(\mu+1/2)}.
\]
Based on Lemma 2 and standard argument for linear regularization theory, see e.g. [18] Theorems 3.1 and 4.1] or [6], we have the following convergence rate results.

**Theorem 4.** Let \( f^k \) be the approximate solution, generated by the scheme (32). Assume that \( \Delta t \in (0, \|K\|) \) is a fixed number. Then, under Assumption 4.

- for \( \mu \in (0, 1/2] \), if the iteration of (32) is terminated according to the a priori stopping rule \( k^* = \delta^{-\frac{1}{2\mu + 1}} \), we have the convergence rate
  \[
  \|f^k - f^*\| = o \left( \delta^{-\frac{2\mu}{2\mu + 1}} \right) \quad \text{as } \delta \to 0.
  \]

- For \( \mu > 1/2 \), if the iteration of (32) is terminated according to the a priori stopping rule \( k^* = \delta^{-\frac{1}{2\mu + 1}} \), we have the convergence rate
  \[
  \|f^k - f^*\| = o \left( \delta^{-\frac{2\mu + 1}{2\mu + 1}} \right) \quad \text{as } \delta \to 0.
  \]

- For general positive \( \mu \), if the iteration of (32) is terminated according to the discrepancy principle (with a fixed positive parameter \( \tau \)), i.e.
  \[
  \|y^\delta - Af^k\| \leq \tau \delta < \|y^\delta - Af^k\|, \quad 0 \leq k < k^*,
  \]
then, it holds that
  \[
  k^* = O \left( \delta^{-\frac{1}{2\mu + 1}} \right), \quad \|f^\delta(k^*) - f^\delta\| = o \left( \delta^{-\frac{1}{2\mu + 1}} \right) \quad \text{as } \delta \to 0.
  \]
Remark 3. By the above theorem, for the newly developed iterative algorithm (32), under Assumption 1, optimal convergence rates are obtained for \( \mu \in (0, 1/2] \) and if the iteration is terminated according to an a priori stopping rule. If \( \mu > 1/2 \) or if the iteration is terminated according to the discrepancy principle, only suboptimal convergence rates can be guaranteed. Nevertheless, the number of iterations for our method (32) is always significantly smaller than the one for Landweber iteration \( \mathcal{O}(\delta^{-2/2}) \).

Remark 4. It is not difficult to show that Theorem 4 also holds for the scheme (32) with \( \Delta t \) replaced by \( \Delta t_k \) such that \( (\Delta t_k)_k \subset [\Delta t_{\min}, \|K\|] \), where \( \Delta t_{\min} > 0 \) is a constant.

Remark 5. It should be noted that the non-symplectic schemes for our second order asymptotical regularization (4) may also provide an accelerated iterative regularization method. For example, consider the following modified Störmer-Verlet scheme (it is not a symplectic method as it belongs to explicit numerical scheme)

\[
\begin{align*}
q^{k+\frac{1}{2}} &= q^k - \frac{\Delta t}{2} (q^k + K^*(y^\delta - Af^k)), \\
f^{k+1} &= f^k + \Delta t q^{k+\frac{1}{2}}, \\
q^{k+1} &= q^{k+\frac{1}{2}} - \frac{\Delta t}{2} q^{k+\frac{1}{2}} + \frac{\Delta t}{2} K^*(y^\delta - Af^{k+1}), \\
f^0 &= f_0, q^0 = 0.
\end{align*}
\] (48)

The above scheme can be written in the form of the three-term semi-iterative method (31) with parameters

\[
a_k = \left(1 - \frac{\Delta t (1 + 2s)}{2t_{k-1}}\right)^2, \quad \omega_k = \frac{\Delta t^2}{2} \left(2 - \frac{\Delta t (1 + 2s)}{2t_{k-1}}\right).
\]

It is not difficult to show that Lemma 2 and hence Theorem 4 also holds for the scheme (48). Consequently, iteration (48) also offers an accelerated iterative regularization method.

4. Application to the diffusion-based bioluminescence tomography (BLT)

4.1. Background of BLT and a reduced mathematical model

In the modern world, biomedical imaging has become extremely important not only for patient care but also for the study of biological structure and function, and for addressing fundamental questions in biomedicine. In molecular imaging, small animal organs and tissues are often labeled with reporter probes that generate detectable signals that can be tracked outside a living body. This technology has been widely used in clinical medicine for investigating tumorigenesis, cancer metastasis, cardiac diseases, etc. In comparison with traditional biomedical imaging approaches such as X-ray computed tomography, positron emission tomography and ultrasound and magnetic resonance imaging, optical molecular imaging has attracted considerable attention for its cost-effectiveness and performance as it directly reveals molecular and cellular activities sensitively [5]. Among various optical molecular imaging techniques, fluorescence molecular tomography [19] and bioluminescence tomography (BLT) [20] are among the most widely used in practice. In contrast with fluorescence imaging, there is no inherent tissue autofluorescence generated by external illumination in bioluminescence imaging, which makes it extremely sensitive. Since bioluminescence imaging cannot provide information about the distribution of an in vivo bioluminescent source, the problem of reconstructing an internal bioluminescent source from the measured bioluminescent signal
on the external surface of a small animal stands an essential mathematical problem in BLT [12].

Bioluminescent photon propagation in biological tissue is governed by the radiative transfer equation (RTE) which has been utilized as the forward model for bioluminescence tomography [17]. However, the RTE is highly dimensional and presents a serious challenge for its accurate numerical simulations given the current level of development in computer software and hardware. Because the mean-free path of the photon is between 500 nm and 1000 nm in biological tissues, which is very small compared to the size of a typical object in this context, the predominant phenomenon in BLT is scattering, which provides a diffusion approximation of the RTE by the following reduced mathematical model [12]

\[
\begin{align*}
-\text{div}(D \nabla u) + \mu_a u &= f \chi_{\Omega_0} \quad \text{in } \Omega, \\
u + 2AD\partial_\nu u &= g^- \quad \text{on } \Gamma,
\end{align*}
\]

where \( u \) denotes the (direction-averaged) photon density, \( D = \left[3(\mu_a + \mu'_s)\right]^{-1} \) with \( \mu_a \) and \( \mu'_s \) being the absorption and reduced scattering coefficients. The boundary \( \Gamma \) of the domain \( \Omega \subset \mathbb{R}^n \) \( (n = 2, 3) \) is assumed to be Lipschitz continuous. \( \partial_\nu \) is the outward normal differentiation operator. \( \Omega_0 \subset \Omega \) is known as a permissible region of the source function, and \( \chi \) is the indicator function such that \( \chi_{\Omega_0}(x) = 1 \) for \( x \in \Omega_0 \), while \( \chi_{\Omega_0}(x) = 0 \), when \( x \not\in \Omega_0 \). \( g^- \) is an incoming flux on \( \Gamma \) and it vanishes when the imaging is implemented in a dark environment. \( A = 1 + \frac{R(x)}{1 - R(x)} \) with \( R(x) \approx -1.4399 \gamma(x)^{-2} + 0.7099 \gamma(x)^{-1} + 0.6681 + 0.0636 \gamma(x) \) and \( \gamma(x) \) being the refractive index of the medium at \( x \in \Gamma \). In the case when \( \Omega \) is a unit circle centered at the origin, \( \mu_a = 0.04, \mu'_s = 1.5 \), and \( A = 3.2 \) with refractive index \( \gamma = 1.3924 \). In BLT, the measurement is the outgoing flux density on the boundary:

\[
g = -D\partial_\nu u \quad \text{on } \Gamma.
\]

If we denote by \( g_1 := g^- + 2Ag \) and \( g_2 := -g \), then the BLT problem (49)-(50) can be formulated as Problem 1. This inverse source problem has been intensively studied in [12, 8, 13, 22, 9, 4, 10, 26] and referenced therein. The essential methodology in these studies is to solve the inverse problem by a two-step strategy. The first step is to adopt Tikhonov variational regularization with \textit{a priori} regularization parameter choice rule to overcome the ill-posedness of original inverse problem, and then solve the regularized PDE-controlled optimization problem by a numerical algorithm (usually we adopt an iterative method). The defects of these existing methods are: (a) Tikhonov regularization exhibits a “strong” saturation phenomenon, i.e., the optimal convergence rate is limited by \( O(\delta^{2/3}) \) with respect to Hölder-type source condition and noise level \( \delta \) of data. (b) \textit{The a priori} stopping rule of regularization parameter is not realistic in practice as it requires some knowledge of the unknown exact solution. (c) Especially for large-scale inverse problems, variational regularization methods are time consuming. In order to overcome these shortcomings, we shall apply the developed accelerated iterative regularization method [32] for the fast solution of Problem 1. It should also be noted that, recently, by assuming the sourcewise representation of source function \( f^\dagger \), the authors in [27] combined the coupled complex boundary method and the expanding compacts method to propose a new regularization method that can calculate a posteriori error estimate efficiently. However, no convergence rate can be derived for such a method.
4.2. Analysis of a mathematical formulation

The aim of this subsection is to reformulate the inverse source problem (12)–(14) as an abstract operator equation so that we can adopt the developed accelerated iterative regularization method in the previous section. We start with the basic assumptions on the system parameter.

**Assumption 2.** \( D \in L^\infty(\Omega) \) and \( D(x) \geq D_0 \) for almost every \( x \in \Omega \); \( \mu_a \in L^2(\Omega) \) and \( \mu_a(x) \geq \mu_0 \) for almost every \( x \in \Omega \). Here, \( D_0 \) and \( \mu_0 \) are two positive constants.

Denote by
\[
V = \{ u : \| u \|_{V,\Omega} < +\infty \}, \quad V_0 = \{ u \in V : u = 0 \ \text{a.e. on} \ \Gamma \},
\]
where
\[
\| u \|_{V,\Omega} = \sqrt{\langle u, u \rangle_{V,\Omega}}, \quad \langle u, v \rangle_{V,\Omega} := (\mu_a u, v)_{L^2(\Omega)} + \langle D\nabla u, \nabla v \rangle_{L^2(\Omega)},
\]
is the weighted \( H^1(\Omega) \) norm. Moreover, we introduce the norms of the trace spaces \( V^{1/2}(\Gamma) \) and \( V^{-1/2}(\Gamma) \) by
\[
\| v \|_{V^{1/2},\Gamma} := \inf_{u \in V} \{ \| u \|_{V,\Omega} : \gamma_0 u = v \}, \quad \| v \|_{V^{-1/2},\Gamma} := \inf_{u \in V} \{ \| u \|_{V,\Omega} : \gamma_1 u = v \},
\]
where \( \gamma_0 : V \to V^{1/2}(\Gamma) \) and \( \gamma_1 : V \to V^{-1/2}(\Gamma) \) are standard trace operators. It is not difficult to show that all of \( V, V_0, V^{1/2}(\Gamma) \) and \( V^{-1/2}(\Gamma) \) are Hilbert spaces, equipped with the corresponding norms (52) and (53), respectively. We remark that if \( D = \mu_a \equiv 1 \), \( V, V^{1/2}(\Gamma) \) and \( V^{-1/2}(\Gamma) \) are reduced to the standard Sobolev spaces \( H^1(\Omega) \), \( H^{1/2}(\Gamma) \) and \( H^{-1/2}(\Gamma) \), respectively. For simplicity, denote \( Q_0 = L^2(\Omega_0), Q = L^2(\Omega) \), and \( Q_\Gamma = L^2(\Gamma) \). Set \( V_{g_1} := \{ v \in V : v = g_1 \ \text{on} \ \Gamma \} \). Define
\[
a(u,v) = \int_\Omega (D\nabla u \cdot \nabla v + \mu_a u v) \, dx, \quad \forall u,v \in V.
\]
Then \( a(\cdot,\cdot) \) is symmetric, continuous and coercive on \( V \). Therefore, by the Lax-Milgram Lemma (71), for any \( f \in Q \), the problems
\[
u_D(f,g_1) \in V_{g_1}, \quad a(u_D(f,g_1),v) = \langle f, v \rangle_{Q_0}, \quad \forall v \in V_0
\]
and
\[
u_N(f,g_2) \in V, \quad a(u_N(f,g_2),v) = \langle f, v \rangle_{Q_0} + \langle g_2, v \rangle_{Q_\Gamma}, \quad \forall v \in V
\]
each have a unique solution. Moreover, a constant \( c > 0 \) exists such that
\[
\| u_D(f,g_1) \|_V \leq c(\| f \|_{Q_0} + \| g_1 \|_{V^{1/2}(\Gamma)}),
\]
\[
\| u_N(f,g_2) \|_V \leq c(\| f \|_{Q_0} + \| g_2 \|_{Q_\Gamma}).
\]
If we define
\[
\left\{
\begin{array}{ll}
u_D(f) = u_D(f,0), & \nu_N(f) = u_N(f,0), \\
u_D(g_1) = u_D(0,g_1), & \nu_N(g_2) = u_N(0,g_2),
\end{array}
\right.
\]
we obtain that \( u_D(f,g_1) = u_D(f) + \nu_D(g_1) \) and \( u_N(f,g_2) = u_N(f) + \nu_N(g_2) \).

Define two operators \( K_D \) and \( K_N \) from \( Q_0 \) to \( H^1(\Omega) \) by
\[
K_D f = u_D(f), \quad K_N q = u_N(q) \quad \forall f \in Q,
\]
and view them as two operators from \( Q_0 \) to \( Q \). Define
\[
K := K_D - K_N, \quad y := \nu_N(g_2) - \nu_D(g_1) \in Q.
\]
It is easy to verify that for any \( f \in Q_0 \),
\[
K f - y = (K_D - K_N) f - y = u_D(f, g_1) - u_N(f, g_2).
\]
Therefore, \( K f = y \) means that \( u_D(f, g_1) = u_N(f, g_2) \).

**Proposition 2.** The operator \( K : Q_0 \to Q \) is compact.

The proof of Proposition 2 can be found in Appendix C. Now, let us consider the case with inexact measurement. Suppose that instead of exact boundary data \( \{g_1, g_2\} \), we are given noisy data \( \{\hat{g}_1^\delta, \hat{g}_2^\delta\} \) satisfying the following assumption.

**Assumption 3.** Let \( g_1, g_1^\delta \in V^{1/2}(\Gamma) \) and \( g_2, g_2^\delta \in V^{-1/2}(\Gamma) \) such that
\[
\|g_1^\delta - g_1\|_{V^{1/2}(\Gamma)} + \|g_2^\delta - g_2\|_{V^{-1/2}(\Gamma)} \leq \mu_0 \delta, \tag{60}
\]
where the noise level \( \delta > 0 \) is known.

At first glance, the assumption \( g_1^\delta \in V^{1/2}(\Gamma) \) looks too strong. However, if \( g_1^\delta \in L^2(\Gamma) \setminus V^{1/2}(\Gamma) \), one can use \( \mathcal{P}_D g_1^\delta \) as the refined Dirichlet data, where \( \mathcal{P}_D : L^2(\Gamma) \to V^{1/2}(\Gamma) \) is the project operator. Since \( V^{1/2}(\Gamma) \) is dense in \( L^2(\Gamma) \), for any \( \varepsilon > 0 \), it holds \( \|\mathcal{P}_D g_1^\delta - g_1^\delta\|_{Q_r} \leq \varepsilon \).

Moreover, as the observed data \( \hat{g}_1^\delta \) is always discrete, one can always interpret the noisy Dirichlet data \( g_1^\delta \in V^{1/2}(\Gamma) \) by using the spline technique for the discrete measurement \( \hat{g}_1^\delta \).

**Proposition 3.** Under Assumption 3, it holds \( \|y^\delta - y\|_Q \leq \delta \), where \( y^\delta = \tilde{u}_N(g_2^\delta) - \tilde{u}_D(g_1^\delta) \).

**Proof.** Define \( v_D := \tilde{u}_D(g_1^\delta) - \tilde{u}_D(g_1) \) and \( v_N := \tilde{u}_N(g_2^\delta) - \tilde{u}_N(g_2) \). Then, \( v_D \) and \( v_N \) satisfy the following BVPs
\[
\begin{aligned}
-\text{div}(D \nabla v_D) &+ \mu_a v_D = 0 \quad \text{in } \Omega, \\
v_D & = g_1^\delta - g_1 \quad \text{on } \Gamma. \tag{61}
\end{aligned}
\]
and
\[
\begin{aligned}
-\text{div}(D \nabla v_N) &+ \mu_a v_N = 0 \quad \text{in } \Omega, \\
\frac{\partial v_N}{\partial n} & = g_2^\delta - g_2 \quad \text{on } \Gamma. \tag{62}
\end{aligned}
\]
Now, let us show that
\[
\|v_D\|_V = \|g_1^\delta - g_1\|_{V^{1/2}(\Gamma)} \quad \text{and} \quad \|v_N\|_V = \|g_2^\delta - g_2\|_{V^{-1/2}(\Gamma)}. \tag{63}
\]

By the definition (53), we have
\[
\|g_2^\delta - g_2\|_{V^{-1/2}(\Gamma)} \leq \|v_D\|_V. \tag{64}
\]
On the other hand, according to equation (61), we have together with (52) that for any \( v \in V \):
\[
\langle v_D, v \rangle_{V, \Omega} = \int_{\Gamma} \frac{\partial v_D}{\partial n}(g_1^\delta - g_1)ds, \]
which implies that
\[
\|v_D\|^2 \leq \left| \int_{\Gamma} \frac{\partial v_D}{\partial n}(g_1^\delta - g_1)ds \right| \leq \|g_1^\delta - g_1\|_{V^{1/2}(\Gamma)} \| \frac{\partial v_D}{\partial n} \|_{V^{-1/2}(\Gamma)}. \]
By the trace theorem, we have \( \| \frac{\partial v_D}{\partial n} \|_{V^{-1/2}(\Gamma)} \leq \|v_D\|_V \). Hence, we derive
\[
\|v_D\|_V \leq \|g_1^\delta - g_1\|_{V^{1/2}(\Gamma)}. \tag{65}
\]
Combine (64) and (65) to obtain the first identity in (63).
Now, consider the second identity in \((63)\). According to \((62)\), for all \(v \in V\), we have together with \((52)\) that
\[
\langle v_N, v \rangle_{V, \Omega} := \langle \mu_a v_N, v \rangle_{L^2(\Omega)} + \langle D \nabla v_N, \nabla v \rangle_{L^2(\Omega)} = \int_\Gamma (g_2^\delta - g_2) \gamma_0 v_N ds. \tag{66}
\]
Set \(v = v_N\) to get \(\|v_N\|_V^2 = \int_\Gamma (g_2^\delta - g_2) \gamma_0 v_N ds\), which gives
\[
\|g_2^\delta - g_2\|_{V^{-1/2}(\Gamma)} = \sup_{\phi \in V^{1/2}(\Gamma)} \frac{\int_\Gamma (g_2^\delta - g_2) \phi ds}{\|\phi\|_{V^{1/2}(\Gamma)}} \geq \phi = \gamma_0 v_N \frac{\int_\Gamma (g_2^\delta - g_2) \gamma_0 v_N ds}{\|\gamma_0 v_N\|_{V^{1/2}(\Gamma)}} = \frac{\|v_N\|_V^2}{\|\gamma_0 v_N\|_{V^{1/2}(\Gamma)}}.
\]
The above inequality together with the trace inequality, i.e. \(\|\gamma_0 v_N\|_{V^{1/2}(\Gamma)} \leq \|v_N\|_V\), gives
\[
\|v_D\|_V \leq \|g_1^\delta - g_1\|_{V^{-1/2}(\Gamma)}. \tag{67}
\]
On the other hand, by using \((63)\) we deduce that
\[
\|g_2^\delta - g_2\|_{V^{-1/2}(\Gamma)} = \sup_{\phi \in V^{1/2}(\Gamma)} \frac{\int_\Gamma (g_2^\delta - g_2) \phi ds}{\|\phi\|_{V^{1/2}(\Gamma)}} = \sup_{\phi \in V^{1/2}(\Gamma)} \frac{\langle v_N, \gamma_0^{-1} \phi \rangle_{V, \Omega}}{||\gamma_0^{-1} \phi||_V} \leq \|v_D\|_V \cdot \sup_{\phi \in V^{1/2}(\Gamma)} \frac{\|\gamma_0^{-1} \phi\|_V}{||\gamma_0^{-1} \phi||_V} \leq \|v_D\|_V,
\]
which implies the second identity of \((63)\) by noting \((67)\).

Finally, by using the definition of \(y^\delta\) and identities \((63)\), we complete the proof by following inequalities
\[
\mu_0 \|y^\delta - y\|_Q \leq \|v_N - v_D\|_V \leq \|g_1^\delta - g_1\|_{V^{1/2}(\Gamma)} + \|g_2^\delta - g_2\|_{V^{-1/2}(\Gamma)} \leq \delta. \tag{68}
\]

Next we discuss the form of \(K^*(Kf - y^\delta)\), which is used in our main algorithms \((32)\) and \((48)\), in the context of the BLT problem. To this end, denote by \(K_D^*\) and \(K_N^*\) the adjoint operators of \(K_D\) and \(K_N\) such that for any \(v \in Q\) and \(f \in Q_0\):
\[
(K_D^* v, f)_{Q_0} = (v, K_D f)_{Q_0}, \quad (K_N^* v, f)_{Q_0} = (v, K_N f)_{Q_0}.
\]
Then \(K^* : Q \to Q_0\) is such that \(K^* = K_D^* - K_N^*\).

For any \(f \in Q_0\), denote by \(u_{DN}(f) = Kf - y^\delta = u_D(f, g_1^\delta) - u_N(f, g_2^\delta)\), and define \(w_D = w_D(u_{DN}(f)) \in V_0\) and \(w_N = w_N(u_{DN}(f)) \in V\) the solutions of the adjoint variational problems
\[
a(v, w_D) = (u_{DN}, v)_Q \quad \forall v \in V_0 \tag{68}
\]
and
\[
a(v, w_N) = (u_{DN}, v)_Q \quad \forall v \in V, \tag{69}
\]
respectively. Then \(K_D^*(Kf - y^\delta) = w_D(u_{DN}(f))|_{\Omega_0}\) and \(K_N^*(Kf - y^\delta) = w_N(u_{DN}(f))|_{\Omega_0}\). Thus, we have
\[
K^*(Kf - y^\delta) = (K_D^* - K_N^*)(Kf - y^\delta) = [w_D(u_{DN}(f)) - w_N(u_{DN}(f))]|_{\Omega_0}. \tag{70}
\]
Similarly, we can give a form of the source condition \([21]\) for our BLT problem. In fact, \([21]\) with \(\mu = 1\) reads that there exists an element \(v_0 \in \mathcal{Q}_0\) such that
\[
f_0 - f^\dagger = (w^*_D - w^*_N)\chi_{\Omega_0},
\]
where \(w^*_D\) and \(w^*_N\) are the solutions of \([68]\) and \([69]\), both with \(u_{DN}\) being replaced by \(u_D(v_0) - u_N(v_0)\), and \(u_D(v_0) = u_D(v_0, 0)\), \(u_N(v_0) = u_N(v_0, 0)\).

We end this section with a source condition for a special case when \(\Omega_0 = \Omega\) and \(\mu = 1/2\). Actually, in this case, \(\mathcal{Q}_0 = \mathcal{Q}\), and consequently, for any \(f, q \in \mathcal{Q}\), we have
\[
(Kf, q)_\mathcal{Q} = \int_\Omega (u_D(f) - u_N(f))qdx = \int_\Omega u_D(f)qdx - \int_\Omega u_N(f)qdx
\]
\[
= \int_\Omega (-\text{div}(D\nabla u_D(q)) + \mu u_D(q))u_D(f)dx
\]
\[
- \int_\Omega (-\text{div}(D\nabla u_N(q)) + \mu u_N(q))u_N(f)dx
\]
\[
= \int_\Omega (-\text{div}(D\nabla u_D(f)) + \mu u_D(f))u_D(q)dx
\]
\[
- \int_\Omega (-\text{div}(D\nabla u_N(f)) + \mu u_N(f))u_N(q)dx
\]
\[
= \int_\Omega (u_D(q) - u_N(q))f_{\text{bdy}} = (f, Kq)_\mathcal{Q},
\]
which means that \(K\) is a self-adjoint operator in \(\mathcal{Q}\). As a result, the source condition with \(\mu = 1/2\) reduces to the existence of an element \(v_0\) and \(\rho \geq 0\) such that \(f_0 - f^\dagger = (K^*K)^{1/2}v_0 = Kv_0 = u_D(v_0) - u_N(v_0)\) and \(\|v_0\| \leq \rho\).

5. Numerical experiments

In this section, we devote ourselves to presenting some numerical examples for demonstrating the effectiveness of the proposed accelerated iterative regularization method \([32]\). We take the diffusion-based bioluminescence tomography considered in Section 4 as example. To that end, with the problem domain \(\Omega\), parameters \(\mu_\alpha, \mu'_\alpha, A,\) Robin data \(g^-,\) and a prescribed true source function \(f^\dagger\), we solve the forward BVP \([49]\) to get \(u^\dagger\). A finite element method of solving \([49]\) is briefly discussed in Appendix D.

The outgoing flux density and the Cauchy data on the boundary are
\[
g = -D\partial_n u^\dagger |_{\Gamma} = \frac{1}{2A}(u^\dagger - g^-), \quad g_1 = g^- + 2Ag, \quad g_2 = -g.
\]

Uniformly distributed noises with the relative error level \(\delta'\) are added to \(g\) to get \(g^\delta\)
\[
g^\delta(x) = [1 + \delta' \cdot (2\text{rand}(x) - 1)]g(x), \quad x \in \Gamma,
\]
where \(\text{rand}(x)\) returns a pseudo-random value drawn from a uniform distribution on \([0, 1]\). The corresponding noisy Cauchy data are \(g_1^\delta = g^- + 2Ag^\delta\) and \(g_2^\delta = -g^\delta\). Then the noise level of the measurement data is calculated by \(\delta = \|g^\delta - y\|_Q\), with \(y = \tilde{u}_N(g_2) - \tilde{u}_D(g_1)\) and \(y^\delta = \tilde{u}_N(g_2^\delta) - \tilde{u}_D(g_1^\delta)\). Without loss of generality, in this section, let \(\mu_\alpha = 0.04, \mu'_\alpha = 1.5, D = 1/[3(\mu_\alpha + \mu'_\alpha)],\) \(A = 3.2\), and \(g^- = 0\), which means the imaging is implemented in a dark environment.
Then, with the noisy data \( g^1 \) and \( g^2 \), properly chosen parameters, e.g. \( s \) and \( \Delta t \), approximate sources \( f^k \) are computed by the proposed accelerated iterative regularization method \((32)\). For the BLT problem, \((32)\) is reduced to

\[
\begin{aligned}
q^{k+1/2} &= q^k - \frac{\Delta t}{t_k} (1 + 2s) q^{k+1/2} - \frac{\Delta t}{2} (w^k_D - w_N^k)\chi_{\Omega_0}, \\
f^{k+1} &= f^k + \Delta t q^{k+1/2}, \\
q^{k+1} &= q^{k+1/2} - \frac{\Delta t}{t_k} (1 + 2s) q^{k+1/2} - \frac{\Delta t}{2} (w^k_{D+1} - w_N^{k+1})\chi_{\Omega_0}, \\
f^0 &= f_0, q^0 = 0,
\end{aligned}
\]

where \( w^k_D \) and \( w_N^k \) are the solutions of \((68)\) and \((69)\), respectively, both with \( u_D(f^k) \) replaced by \( u_D(f^k) \). \( w^k_{D+1} \) and \( w_N^{k+1} \) have similar definitions. In the following, for the conciseness of the statements, we only consider the case that using Morozov’s discrepancy principle \((16)\) to control the iterative procedure, namely that the iteration stops when \( \|y^{\delta} - Af^k\| = \|u_D(f^k, g^1) - u_N(f^k, g^2)\| \leq \tau \delta \), where \( u_D(f^k, g^1) \) and \( u_N(f^k, g^2) \) are the solutions of \((55)\) and \((56)\), both with \( q \) being replaced by \( f^k \), and with \( g_1 \) and \( g_2 \) being replaced by \( g^1 \) and \( g^2 \), respectively. Moreover, the initial guess \( f_0 \) is chosen so that the condition of Lemma \((1)\) is satisfied: \( \|y^{\delta} - A f_0\| = \|u_D(f_0, g^1) - u_N(f_0, g^2)\| > \tau \delta \), where \( u_D(f_0, g^1) \) and \( u_N(f_0, g^2) \) have similar definitions as \( u_D(f^k, g^1) \) and \( u_N(f^k, g^2) \) above.

We use \( N_{\text{max}} := 5000 \) as the maximal number of iterations where the iteration \((72)\) stops, which may have different values in different experiments. To assess the accuracy of the approximate solutions, we define the \( L^2 \)-norm relative error for an approximate solution \( f^k \): 

\[
E_k := \|f^k - f^0\|_{0, \Omega} / \|f^0\|_{0, \Omega}.
\]

All experiments in Subsection \(5.1\)–\(5.2\) are implemented for the following two examples:

**Example 1**: \( \Omega := \{(x_1, x_2) \in \mathbb{R}^2 | x_1^2 + x_2^2 < 1\} \), \( f^1(x_1, x_2) = (1 + x_1 + x_2)\chi_{\Omega_0} \) with \( \Omega_0 := \{(x_1, x_2) \in \mathbb{R}^2 | -0.5 < x_1, x_2 < 0.5\} \). The measurements are computed on a mesh with mesh size \( h = 0.01386, 144929 \) nodes and 288768 elements.

**Example 2**: \( \Omega \) is the same as Example 1, \( f^1(x_1, x_2) = (1 + x_1 + x_2)\chi_{\Omega_1} + e^{1+x_1+x_2} \chi_{\Omega_2} \) with \( \Omega_1 := \{(x_1, x_2) \in \mathbb{R}^2 | (x_1 + 0.5)^2 + x_2^2 < 0.01\} \) and \( \Omega_2 := \{(x_1, x_2) \in \mathbb{R}^2 | (x_1 - 0.5)^2 + x_2^2 < 0.01\} \). The measurements are computed on a mesh with \( h = 0.01228, 156225 \) nodes and 311296 elements.

For Example 1, all approximate sources are reconstructed over a mesh with mesh size \( h = 0.0744, 2325 \) nodes and 4512 elements. For Example 2, all approximate sources are reconstructed over a mesh with mesh size \( h = 0.0678, 2505 \) nodes and 4864 elements.

### 5.1. Influence of parameters

The purpose of this subsection is to explore the dependence of the solution accuracy and the convergence speed on \( \tau > 0 \), time step size \( \Delta t \), parameter \( s \) in the damping parameter \( (1 + 2s)/t \), and thus to give a guide on the choices for them in practice. For focusing on the effect of these parameters on the iteration \((72)\), we fix \( \delta' = 5\% \) in this subsection. Moreover, in the remaining part of this section, we simply set \( f_0 = 1, g_0 = 0 \).

We first investigate the influence of parameter \( \tau \) on the convergence rate. For this purpose, we additionally set \( s = 1 \), and \( \Delta t = 0.125 \) for Example 1 and \( \Delta t = 0.25 \) for Example 2. The detailed iterative numbers \( k^* \) and the corresponding \( L^2 \)-norm relative errors \( E_{k^*'} \) for different values of \( \tau \) are shown in Table \((1)\), which shows that the smaller \( \tau \) is, the higher the iterative number for stopping \((72)\) is. It is no surprise that the parameter \( \tau \) does not involve the
Table 1. The iterative number \( k^* \) and the corresponding relative error \( E_{k^*} \) vs \( \tau \).

| \( \tau \)   | Example 1 | Example 2 |
|-------------|-----------|-----------|
|             | \( E_{k^*} \) | \( k^* \) | \( E_{k^*} \) | \( k^* \) |
| 2^{-5}      | 1.4374e-1 | \( N_{\text{max}} \) | 4.1898e-2 | \( N_{\text{max}} \) |
| 2^{-4}      | 1.4374e-1 | \( N_{\text{max}} \) | 4.1898e-2 | \( N_{\text{max}} \) |
| 2^{-3}      | 1.4374e-1 | \( N_{\text{max}} \) | 4.1898e-2 | \( N_{\text{max}} \) |
| 2^{-2}      | 1.4374e-1 | \( N_{\text{max}} \) | 4.1898e-2 | \( N_{\text{max}} \) |
| 2^{-1}      | 1.4374e-1 | \( N_{\text{max}} \) | 3.7214e-2 | 615 |
| 1           | 2.2069e-3 | 77        | 6.753e-2  | 415 |
| 2           | 5.0111e-2 | 66        | 9.3589e-2 | 124 |
| 2^2         | 1.1240e-1 | 55        | 1.1617e-1 | 44  |
| 2^3         | 2.3084e-1 | 36        | 1.2571e-1 | 41  |
| 2^5         | 3.7765e-1 | 1         | 1.4729e-1 | 38  |

computation of the approximate solutions itself. It is used in the stop criterion and affects
only the iterative number at which the iteration (72) stops. In the remaining experiments, we
choose \( \tau = 1 \) for Example 1 and \( \tau = 5 \) for Example 2.

Now we investigate the influence of time step size \( \Delta t \) on the solution accuracy and the
convergence rate. To this end, set \( s = 1 \) and \( \tau = 1 \) for Example 1, \( \tau = 5 \) for Example 2. The
iterative numbers \( k^* \) and the corresponding L2-norm relative errors \( E_{k^*} \) are given in Table
2, which shows that the bigger the time step size \( \Delta t \) is, the faster the iteration is. However,
our experiments suggest that \( \Delta t \) should not be too big, e.g. \( \Delta t \leq 0.125 \). Otherwise, the
iteration will blow up as it breaks the consistency of the numerical scheme. In the remaining
experiments, we choose \( \Delta t = 0.125 \).

Table 2. The iterative number \( k^* \) and the corresponding relative error \( E_{k^*} \) vs \( \Delta t \).

| \( \tau \)   | Example 1 | Example 2 |
|-------------|-----------|-----------|
|             | \( E_{k^*} \) | \( k^* \) | \( E_{k^*} \) | \( k^* \) |
| 2^{-10}     | 2.1189e-1 | \( N_{\text{max}} \) | 4.2264e-1 | \( N_{\text{max}} \) |
| 2^{-9}      | 2.0730e-3 | 4910      | 1.3920e-1 | \( N_{\text{max}} \) |
| 2^{-8}      | 2.0730e-3 | 2455      | 1.154e-1  | 2898 |
| 2^{-7}      | 2.0572e-3 | 1288      | 1.1545e-1 | 1449 |
| 2^{-6}      | 2.0572e-3 | 614       | 1.154e-1  | 724  |
| 2^{-5}      | 2.0571e-3 | 307       | 1.0686e-1 | 671  |
| 2^{-4}      | 2.2017e-3 | 154       | 1.0587e-1 | 341  |
| 2^{-3}      | 2.2069e-3 | 77        | 1.0476e-1 | 177  |
| 2^{-2}      | Divergence | -        | 1.0674e-1 | 84   |
| 2^{-1}      | Divergence | -        | Divergence | -    |

Finally, we discuss the influence of \( s \) in damping parameter \((1 + 2s)/t\) on the solution
accuracy and the convergence rate. In the experiments, set \( \tau = 1, \Delta t = 0.125 \) for Example
1 and \( \tau = 5, \Delta t = 0.25 \) for Example 2. The required number of iterations \( k^* \) and the
corresponding relative error \( E_{k^*} \) for different values of \( s \) are given in Table 3, which indicates
that in general there is no optimal choice of parameter \( s \). For both model problems, \( s > -0.1 \)
is sufficient to produce satisfactory solutions. However, small values of \( s \), i.e. \( s < 1 \), usually
Table 3. The iterative number \( k^* \) and the corresponding relative error \( E_{k^*} \) vs \( s \).

| \( s \)     | \( E_{k^*} \) | \( k^* \) | \( E_{k^*} \) | \( k^* \) |
|------------|----------------|---------|----------------|---------|
| -0.499     | 1.3357e-2      | 36      | 7.7630e-1      | \( N_{\text{max}} \) |
| -0.4       | 8.3658e-3      | 38      | 3.1754e-1      | 3497    |
| -0.3       | 2.7977e-3      | 40      | 2.6627e-1      | 3417    |
| -0.2       | 7.8878e-3      | 43      | 2.3198e-1      | 3377    |
| -0.1       | 1.9813e-3      | 45      | 7.2336e-2      | 1038    |
| 0          | 3.2890e-3      | 48      | 6.0992e-2      | 958     |
| \( 2^{-5} \) | 4.2415e-3      | 49      | 6.0058e-3      | 918     |
| \( 2^{-4} \) | 5.1298e-3      | 50      | 5.8946e-2      | 878     |
| \( 2^{-3} \) | 6.6319e-3      | 52      | 5.7770e-2      | 838     |
| \( 2^{-2} \) | 2.2341e-3      | 55      | 5.4845e-2      | 441     |
| \( 2^{-1} \) | 5.0426e-3      | 63      | 6.3631e-2      | 241     |
| 1          | 2.2069e-3      | 77      | 9.5095e-2      | 121     |
| 2          | 2.1025e-3      | 104     | 1.1467e-1      | 61      |
| \( 2^2 \)  | 2.2495e-3      | 152     | 1.1357e-1      | 89      |
| \( 2^3 \)  | 2.2862e-3      | 226     | 1.1240e-1      | 133     |
| \( 2^4 \)  | 2.3592e-3      | 331     | 1.1161e-1      | 195     |
| \( 2^5 \)  | 2.4120e-3      | 477     | 1.1111e-1      | 282     |

bring the oscillation in solution accuracy during iterations (cf. Figure 1), and fail to offer a better result in many cases (e.g. Example 2 in our experiments). Furthermore, we conclude from Table 3 that for large values of \( s \), i.e. \( s \geq 1 \), the factor \( s \) has less effect on the solution accuracy, but has considerable influence on the iterative number. We remark that for large values of \( s \), one can use a small \( \tau \) to improve the solution accuracy. For instance, for \( s = 2^5 \), by using \( \tau = 1 \) one can produce an approximate solution with \( E_{k^*} = 5.6243 \times 10^{-2} \) which is significantly smaller than \( E_{k^*} = 1.1111 \times 10^{-1} \) corresponding to \( \tau = 5 \). It is suggested that a value of \( s \) near 1 produces satisfactory results in both solution accuracy and the iterative number for Examples 1 and 2, which coincides with the empirical results about the optimal parameter choice for the Nesterov’s method. Therefore, in the remaining experiments, we set \( s = 1 \).

5.2. Comparison with other methods

In this subsection, we compare the behaviors regarding the solution accuracy and the convergence rate between scheme (72), the modified Störmer-Verlet scheme (18), two well-known acceleration methods: the Nesterov’s method, the \( \nu \)-method, and the Landweber method (2). For the BLT: Problem 1, the modified Störmer-Verlet scheme (18) has the form

\[
\begin{align*}
q^{k+1/2} &= q^k - \frac{\Delta t}{2} \frac{1+2s}{t_k} q^k - \frac{\Delta t}{2} \left( w_D^k - w_N^k \right) \chi_{\Omega_0}, \\
f^{k+1} &= f^k + \Delta t q^{k+1/2}, \\
q^{k+1} &= q^{k+1/2} - \frac{\Delta t}{2} \frac{1+2s}{t_{k+1}} q^{k+1/2} - \frac{\Delta t}{2} \left( w_D^{k+1} - w_N^{k+1} \right) \chi_{\Omega_0}, \\
f^0 &= f_0, q^0 = 0.
\end{align*}
\] (73)

In our simulations, for the BLT problem, the \( \nu \)-method is defined by

\[
\begin{align*}
f^k &= f^{k-1} + \mu_k (f^{k-1} - f^{k-2}) - \omega \cdot \omega_k \cdot (w_D^{k-1} - w_N^{k-1}) \chi_{\Omega_0}, \\
f^0 &= f^{-1} = f_0
\end{align*}
\] (74)
Figure 1. Evolutions of $E_k$ vs. $k$ for different values of $s$ (Example 2).

with $\mu_1 = 0$, $\omega_1 = (4\nu + 2)/(4\nu + 1)$ and

$$
\mu_k = \frac{(k - 1)(2k - 3)(2k + 2\nu - 1)}{(k + 2\nu - 1)(2k + 4\nu - 1)(2k + 2\nu - 3)}.
$$
\[ \omega_k = 4\frac{(2k + 2\nu - 1)(k + \nu - 1)}{(k + 2\nu - 1)(2k + 4\nu - 1)} \text{ for } k > 1, \]

where \( \omega > 0 \) is the weight. We select the Chebyshev method as our special \( \nu \)-method, i.e., \( \nu = 1/2 \). Note that \( \omega_{\text{norm}} = 1 \) in the convolutional \( \nu \)-method (cf. \[6 \] § 6.3]) as it works for a normalized operator equation \( Kf = y \) with \( \|K\| \leq 1 \). For our BLT problem, \( \omega \) in (74) plays the role of normalization, and it can be set as \( \omega = \omega_{\text{norm}}(:= 1/\|K^*K\|) \), which can be calculated by

\[ \omega_{\text{norm}} = \frac{\|1\|_{Q_0}}{\|(w_D(1, g_1^d) - w_N(1, g_2^d)) - (w_D(0, g_1^d) - w_N(0, g_2^d))\|_{Q_0}}. \]

The Nesterov’s method is defined by (18)

\[
\begin{align*}
  z_k &= f^k + \frac{k-1}{k+\alpha-1} (f^k - f^{k-1}), \\
  f^{k+1} &= z_k - \omega(w_D^k - w_N^k)\chi_{\Omega_0}, \\
  f^0 &= f^{-1} = f_0
\end{align*}
\] (75)

with \( 0 < \alpha \geq 3 \) and \( 0 < \omega < \omega_{\text{norm}} \). In all simulations, we choose \( \alpha = 3 \). For Example 1, \( \omega_{\text{norm}} \approx 0.00542264152263 \), we set \( \omega = 0.005 \); for Example 2, \( \omega_{\text{norm}} \approx 0.021370788062004 \), we set \( \omega = 0.02 \).

The Landweber method (2) has the form

\[
\begin{align*}
  f^{k+1} &= f^k - \Delta t(w_D^k - w_N^k)\chi_{\Omega_0}, \\
  f^0 &= f_0
\end{align*}
\] (76)

with \( 0 < \Delta t < 2\omega_{\text{norm}} \). For Example 1, we set \( \Delta t = 2 \times 0.005 = 0.01 \); for Example 2, we set \( \Delta t = 2 \times 0.02 = 0.04 \).

As suggested by Subsection 5.1 in our methods (72) (termed as “ARM”) and (73) (termed as “MSVM”), we set \( \Delta t = 0.125, s = 1 \) for Example 1 \( \Delta t = 0.25, s = 1 \) for Example 2. In all methods, the initial guess \( f_0 = 0 \) and the iterations stop when \( \|g^\delta - Kf^k\| = \|u_D(f^k, g_1^d) - u_N(f^k, g_2^d)\| < \tau \delta \) with \( \tau = 1.2 > 1 \) for Example 1 and \( \tau = 5 > 1 \) for Example 2. We note that when the noisy level is large, bigger \( \tau \) is suggested so that iterations can stop before the solution accuracy gets worse.

The results of the simulations are presented in Table 4, from which we conclude that, with properly chosen parameters, all the above mentioned methods are stable and can produce satisfactory solutions. Moreover, on the one hand, compared with the conventional Landweber method, all of the other methods produce better accuracy with considerably fewer iterations; on the other hand, compared with the well-known accelerated regularization methods, i.e. the \( \nu \)-method and the Nesterov’s method, the proposed ARM and MSVM converge even faster for the given model problems.

We note that inverse source problems with only one measurement on the boundary do not have a unique solution. In the context of the BLT problem, one cannot distinguish between a strong source over a small region and a weak source over a large region. Therefore, in all the above experiments, we are assumed to know exactly the positions of sources (i.e. the geometry of \( \Omega_0 \)). We can also apply the proposed method to other linear inverse problems and compare the behavior of different methods. Another well-known linear inverse problem is the Cauchy problem of finding \( (\phi, t) \) on unaccessible boundary \( \Gamma_u \) from the Cauchy data.
(Φ, T) on accessible boundary Γ_a such that the following relations hold:

\[
\begin{align*}
-\Delta u &= 0 \quad \text{in } \Omega, \\
\partial_\nu u &= \Phi, \quad u = T \quad \text{on } \Gamma_a, \\
\partial_\nu u &= \phi, \quad u = t \quad \text{on } \Gamma_u.
\end{align*}
\]

In contrast to the BLT problem, the Cauchy problem above admits solution uniqueness, provided a solution exists. We can expect good behavior of the proposed method for this Cauchy problem. However, for the conciseness of the paper, we omit these numerical results.

6. Conclusions

In this paper we have proposed a new class of accelerated regularization methods for solving ill-posed linear operator equations. A series of theoretical results including limiting behavior and convergence rates are proved. Moreover, as an application of the proposed method, in this paper, a model problem arising from bioluminescence tomography is discussed in detail. Since the proposed methods are comparable to the Nesterov’s acceleration method and the ν-method about the convergence rate and the solution accuracy, they are promising approaches which merit further theoretical and numerical development as well as more extensive comparison to state-of-the-art methods. Similar to Nesterov’s acceleration method, the introduced iterative regularization methods can also be used to solve to some non-linear ill-posed problems. However, for performing a rigorous theoretical analysis, the concept of acceleration in the sense of regularization theory should be extended so that it can be used for evaluating general non-linear regularization methods.

Acknowledgement

The work of R. Gong is supported by the Natural Science Foundation of China (No. 11401304) and the Fundamental Research Funds for the Central Universities (No. NS2018047). The work...
of Y. Zhang is supported by the Alexander von Humboldt foundation through a postdoctoral researcher fellowship, and the work of B. Hofmann is supported by the German Research Foundation (DFG-grant HO 1454/12-1).

Appendix A: Proof of Proposition \[1\]
For simplicity, we only consider the case of positive \(s\). For the case \(s \in (-1/2, 0)\), we refer to the similar result, presented in [3, Lemma 6.1]. The general solution to (8) is

\[
\xi_j(t) = \lambda_j^{-1} \langle y^\delta, v_j \rangle + \left\{ \begin{array}{ll}
\frac{C_{1,s}^j}{(\lambda_j)^s} J_s(\lambda_j t) + \frac{C_{2,s}^j}{(\lambda_j)^s} J_s(\lambda_j t), & \text{if } s \in \mathbb{N} \cup \{0\}, \\
\frac{C_{1,s}^j}{(\lambda_j)^s} J_s(\lambda_j t) + \frac{C_{2,s}^j}{(\lambda_j)^s} J_s(-\lambda_j t), & \text{if } s \notin \mathbb{N},
\end{array} \right.
\]

where \(Y_s\) denotes the Bessel functions of second kind. In order to determine the constants \(C_{1,s}^j\) and \(C_{2,s}^j\) from the initial conditions, we distinguish three different cases: (i) \(s = 0\), (ii) \(s \in \mathbb{N}\), and (iii) \(s \notin \mathbb{N}\). We show that for all of three cases \(C_{2,s}^j = 0\) according to the boundedness of initial data. In case (i), by using the divergence behaviour \(Y_0(\lambda_j t) = \mathcal{O}(\log(\lambda_j t))\) as \(t \to 0\) [1, (9.1.12)], \(C_{2,0}^j\) must be zero. For case (ii), the asymptotic, cf. [11, (9.1.11)], \(Y_s(\lambda_j t) = \mathcal{O}((\lambda_j t)^{-s})\) as \(t \to 0\) implies \(\frac{Y_s(\lambda_j t)}{(\lambda_j t)^s} = \mathcal{O}((\lambda_j t)^{-2s})\). Therefore, \(C_{2,s}^j = 0\) for all \(s \in \mathbb{N}\). Now, consider the last case. According to the asymptotic \(J_{-s}(\lambda_j t) = \mathcal{O}((\lambda_j t)^{-s})\) as \(t \to 0\), cf. [11, (9.1.10)], \(C_{2,s}^j = 0\) for all \(s \notin \mathbb{N}\).

By the above analysis, the general solution to (8) bounded initial data should be

\[
\xi_j(t) = \frac{C_{1,s}^j}{(\lambda_j)^s} J_s(\lambda_j t) + \lambda_j^{-1} \langle y^\delta, v_j \rangle.
\]

By the initial data \(f_0 = \sum_j \langle f_0, u_j \rangle u_j\) and the limit \(\lim_{t \to 0^+} \frac{J_s(\lambda_j t)}{(\lambda_j t)^s} = \frac{1}{2^{s+1}}\), we conclude that

\[
C_{1,s}^j = 2^s \Gamma(s + 1) \left( \langle f_0, u_j \rangle - \lambda_j^{-1} \langle y^\delta, v_j \rangle \right),
\]

which gives the desired formula for \(\xi_j(t)\).

Finally, check that \(\dot{\xi}_j(0) = 0\). It can be done by the following limit

\[
\dot{\xi}_j(0) = \lim_{t \to 0^+} \frac{1 - 2^s \Gamma(s + 1) \frac{J_s(\lambda_j t)}{(\lambda_j t)^s} \left( \lambda_j^{-1} \langle y^\delta, v_j \rangle - \langle f_0, u_j \rangle \right)}{t} = 0
\]

by noting that [11, (9.1.10)]

\[
\frac{J_s(\lambda_j t)}{(\lambda_j t)^s} = \frac{1}{2^s \Gamma(s + 1)} + \mathcal{O}((\lambda_j t)^2)\text{ as } t \to 0.
\]

Appendix B: Proof of Lemma \[1\]
This proof uses the technique in [2]. Consider the Lyapunov function of (11) by \(\mathcal{E}(t) = \frac{1}{2} \| \dot{f}^\delta(t) \|^2 + \| K f^\delta(t) - y^\delta \|^2 \). It is easy to show that

\[
\dot{\mathcal{E}}(t) = -\frac{1 + 2s}{t} \| \dot{f}^\delta(t) \|^2 \leq 0.
\]

Hence, \(\mathcal{E}(t)\) is non-increasing, and \(\mathcal{E}(\infty) := \lim_{t \to \infty} \mathcal{E}(t)\) exists by noting that \(\mathcal{E}(t) \geq 0\) for all \(t\). Now, consider the function \(\epsilon(t) = \frac{1}{2} \| \dot{f}^\delta(t) - f^\delta \|^2\). It is not difficult to obtain

\[
\dot{\epsilon}(t) + \frac{1 + 2s}{t} \epsilon(t) + \| K f^\delta(t) - y^\delta \|^2 \leq \| \dot{f}^\delta(t) \|^2.
\]
Divide this expression by $t$ to obtain
\[
\frac{1}{t} \dot{e}(t) + \frac{1 + 2s}{t^2} \dot{e}(t) + \frac{1}{t} \mathcal{E}(t) \leq \frac{3}{2t} \| f\delta(t) \|^2.
\]
Integrating the above inequality from 1 to $t$ and using integration by parts for $\dot{e}(t)$, we obtain
\[
\int_1^t \frac{\mathcal{E}(\tau)}{\tau} d\tau \leq \dot{e}(1) - \frac{\dot{e}(t)}{t} - 2(1 + s) \int_1^t \frac{\dot{e}(\tau)}{\tau^2} d\tau + \frac{3}{2} \int_1^t \frac{\| f\delta(\tau) \|^2}{\tau} d\tau.
\]  
(81)
On the one hand, using the integration by parts and the positivity of functional $e(\cdot)$, we have
\[
\int_1^t \frac{\dot{e}(\tau)}{\tau^2} d\tau = \frac{e(t)}{t^2} - e(1) + 2 \int_1^t \frac{e(\tau)}{\tau^3} d\tau \geq -e(1).
\]  
(82)
On the other hand, relation (79) gives
\[
\int_1^t \frac{\| f\delta(\tau) \|^2}{\tau} d\tau = \frac{\mathcal{E}(1) - \mathcal{E}(t)}{1 + 2s}.
\]  
(83)
Combine (81)–(83) to get
\[
\int_1^t \frac{\mathcal{E}(\tau)}{\tau} d\tau \leq \dot{e}(1) - \frac{\dot{e}(t)}{t} + 2(s + 1)e(1) + \frac{3(\mathcal{E}(1) - \mathcal{E}(t))}{2(1 + 2s)} - \frac{3\mathcal{E}(t)}{2(1 + 2s)}.
\]  
(84)
where $C(1) = \dot{e}(1) + 2(s + 1)e(1) + \frac{3\mathcal{E}(1)}{2(1 + 2s)}$ collects the constant terms. Therefore, for any $T \geq t > 1$, we have
\[
\mathcal{E}(T) \int_1^t \frac{1}{\tau} d\tau + \frac{3\mathcal{E}(T)}{2(1 + 2s)} \leq C(1) - \frac{\dot{e}(t)}{t}.
\]  
(86)
by noting the non-increasing of Lyapunov function $\mathcal{E}(t)$. Rewrite (86) as $\mathcal{E}(T) \left[ \ln(t) + \frac{3}{2(1 + 2s)} \right] \leq C(1) - \frac{\dot{e}(t)}{t}$, and then integrate it from $t = 1$ to $t = T$ to derive
\[
\mathcal{E}(T) \left( T \ln(T) + 1 - T + \frac{3}{2(1 + 2s)}(T - 1) \right) \leq C(1)(T - 1) - \int_1^T \frac{\dot{e}(t)}{t} dt.
\]  
(87)
Moreover, using the integration by parts and the positivity of functional $e(\cdot)$, we have
\[
\int_1^T \frac{\dot{e}(\tau)}{\tau} d\tau = \frac{e(T)}{T} - e(1) + \int_1^T \frac{e(t)}{t^2} dt \geq -e(1).
\]  
(88)
By combining (87) and (88), we deduce that
\[
\mathcal{E}(T) \left( T \ln(T) + C_1T + C_2 \right) \leq C(1)T + C_3,
\]  
(89)
where $C_1 = \frac{3}{2(1 + 2s)} - 1$, $C_2 = -C_1$ and $C_3 = e(1) - C(1)$ are three constants.

Inequality (89) immediately yields $\mathcal{E}(\infty) \leq 0$. By the non-negativity of Lyapunov function $\mathcal{E}(\cdot)$, we conclude
\[
\mathcal{E}(\infty) = 0.
\]  
(90)

The continuity of $\chi(T)$ is obvious as our problem is linear. Hence, from (90) and the assumption of the lemma, we conclude that
\[
\lim_{T \to \infty} \chi(T) \leq (1 - \tau)\delta < 0 \quad \text{and} \quad \chi(0) = \| Kf_0 - y \delta \| - \tau\delta > 0,
\]  
which implies the existence of the root of $\chi(T)$.
Accelerated regularization methods

Appendix C: Proof of Proposition 2

Let \( \{f^n\}_n \subset Q_0 \) be bounded. Then there is a subsequence, denoted again by \( \{f^n\}_n \), which converges weakly in \( Q_0 \) to some element \( f^* \in Q_0 \) because of the reflexivity of space \( Q_0 \). Let \( u^n_D = u_D(f^n) \), \( u^n_N = u_N(f^n) \), i.e., \( u^n_D \in V_0 \), \( u^n_N \in V \), and

\[
a(u^n_D, v) = \langle f^n, v \rangle_{Q_0} \quad \forall \, v \in V_0, \tag{91}
\]

\[
a(u^n_N, v) = \langle f^n, v \rangle_{Q_0} \quad \forall \, v \in V. \tag{92}
\]

Then \( \{u^n_D\}_n \) and \( \{u^n_N\}_n \) are bounded in \( V \) from the properties (57) and (58). Hence, we can extract two further subsequence, denoted again by \( \{u^n_D\}_n \) and \( \{u^n_N\}_n \), which converge weakly in \( V \) and strongly in \( Q \) to \( u^*_D \in V_0 \) and \( u^*_N \in V \), respectively. Let \( n \to \infty \) in (91) and (92) to get \( u^*_D = u_D(f^*) \) and \( u^*_N = u_N(f^*) \). Strong convergence of \( \{u^n_D\}_n \) to \( u^*_D \) in \( V \) follows from

\[
\mu_0 \|u^n_D - u^*_D\|^2_V \leq a(u^n_D - u^*_D, u^n_D - u^*_D) = \int_{Q_0} (f^n - f^*)(u^n_D - u^*_D) \, dx \to 0
\]
as \( n \to \infty \). Similarly, \( u^*_N \to u^*_N \) as \( n \to \infty \).

Denote \( g^n = K f^n \). Then \( \{g^n\}_n \) is bounded in \( Q \). Repeating the above argument, we conclude that there exists an element \( f^* \in Q \) such that

\[
g^n \to g^* \text{ in } Q, \quad u_D(g^n) \to u_D(g^*), \quad u_N(g^n) \to u_N(g^*) \text{ in } V \text{ as } n \to \infty.
\]

Therefore, \( \forall \, v \in Q \),

\[
(K f^* - g^*, v)_Q = \lim_{n \to \infty} (K f^n - g^n, v)_Q = \lim_{n \to \infty} (u_D(f^n) - u_N(f^n) - g^n, v)_Q
\]

\[
= \lim_{n \to \infty} (u_D(f^n) - u_N(f^n) - g^*, v)_Q = \lim_{n \to \infty} (K f^n - g^n, v)_Q = 0.
\]

Thus, we have \( g^* = K f^* \). Consequently, strong convergence of \( g^n \) to \( g^* \) in \( Q \) follows from

\[
\|g^n - g^*\|^2_Q = \|K f^n - K f^*\|^2_Q = \|u_D(f^n) - u_D(f^*) - (u_N(f^n) - u_N(f^*))\|^2_Q
\]

\[
\leq 2\|u_D(f^n) - u_D(f^*)\|^2_Q + 2\|u_N(f^n) - u_N(f^*)\|^2_Q \to 0
\]
as \( n \to \infty \), and the proof is completed.

Appendix D: Finite element discretization of boundary value problems

In this appendix, we discuss the numerical implementations of (55) and (56) by standard finite element method. We use linear finite element space for an approximation of the light source space \( Q_0 \). Specifically, let \( \{T_{0,H}\}_H \) be a regular family of triangulations over domains \( \overline{\Omega}_0 \subset \overline{\Omega} \) with meshsize \( H > 0 \). For each triangulation \( T_{0,H} = \{K_H\} \), define finite element space \( Q_{0,H}^H = \{q \in C(\overline{\Omega}_0) \mid q|_{K_H} \in \mathcal{P}_1(K), \quad \forall \, K_H \in T_{0,H}\} \), where \( \mathcal{P}_k \) represents the space of all polynomials of degree no greater than \( k \). Let \( \{T_h\}_h \) be a regular family of triangulations over domains \( \overline{\Omega} \subset \mathbb{R}^d \) with a mesh size \( h > 0 \). For each triangulation \( T_h = \{K_h\} \), define finite element spaces \( V_h \) and \( V_0^h \) as follows.

\[
V_h := \{v \in C(\overline{\Omega}) \mid v|_{K_h} \in \mathcal{P}_1, \quad \forall \, K_h \in T_h\}, \quad V_0^h = V_h \cap V_0.
\]

Moreover, we use the symbol \( g_0^\delta + V_0^h \) for the set

\[
\{v \in V_h \mid v(x_i) = g_0^\delta(x_i) \quad \forall \text{ vertex } x_i \in K_h \cap \Gamma, \quad \forall \, K_h \in T_h\}.
\]
For each $f \in Q_0$, the finite element discretization of (63) and (69) read

$$u_D^h := u_D^h(f, g_1^\delta) \in \mathcal{V}_0^h, \quad a(u_D^h, v) = \langle f, v \rangle_{Q_0} \quad \forall \, v \in \mathcal{V}_0^h,$$  

$$u_N^h := u_N^h(f, g_2^\delta) \in \mathcal{V}^h, \quad a(u_N^h, v) = \langle f, v \rangle_{Q_0} + \langle g_2^\delta, v \rangle_{Q_T} \quad \forall \, v \in \mathcal{V}^h. \quad (94)$$

Similar to the continuous case, we use the symbols $u_D^h(f)$, $u_N^h(q)$ and $\tilde{u}_D^h(g_1^\delta)$ for $u_D^h(f, 0)$, $\tilde{u}_D^h(0, g_1^\delta)$, $u_N^h(q, 0)$ and $u_N^h(0, g_2^\delta)$, respectively.

Suppose that $\mathcal{T}_{0,H}$ and $\mathcal{T}_h$ are consistent, i.e., the triangulation $\mathcal{T}_h$ on $\overline{\Omega}_0$, and let $n_0$ and $n$ be the numbers of nodes of the triangulations $\mathcal{T}_{0,H}$ and $\mathcal{T}_h$. Denote $\varphi_i(x) \in \mathcal{V}^h$, $1 \leq i \leq n$, be the node basis functions of the finite element space $\mathcal{V}^h$ associated with grid nodes $x_i \in \overline{\Omega}$. Let $x_{i,j} \in \Omega_{0,i}, 1 \leq j \leq n_0$ be the nodes of $\mathcal{T}_{0,H}$, and $\varphi_i(x) \in \mathcal{V}^h$ the corresponding basis functions. Then, the approximate source function $f^H$ of $f$ can be expressed by $f^H = \sum_{j=1}^{n} f^H_j \varphi_{i,j}$ with $f^H_j = f(x_{i,j})$. For the problems (93) and (94), the solutions $u_D^h \in \mathcal{V}_0^h$ and $u_N^h \in \mathcal{V}^h$ can be expanded by $u_D^h = \sum_{i=1}^{n} u_{D,i} \varphi_i$ and $u_N^h = \sum_{i=1}^{n} u_{N,i} \varphi_i$, respectively, where $u_{D,i} = u_D^h(x_{i})$ and $u_{N,i} = u_N^h(x_{i})$.

Denote $I = \{1, 2, \cdots, n\}$, $I_0 = \{1, 2, \cdots, n_0\}$, $I_b = \{i \in I | x_i \in \Gamma\}$, and define

$$S = (s_{i,j}), \quad s_{i,j} = \int_{\Omega} D \nabla \varphi_i \nabla \varphi_j \, dx, \quad i, j \in I,$$

$$M = (m_{i,j}), \quad m_{i,j} = \int_{\Omega} \mu_a \varphi_i \varphi_j \, dx, \quad i, j \in I,$$

$$M_0 = (m_{0,j}), \quad m_{0,j} = \int_{\Omega} \mu_0 \varphi_i \varphi_j \, dx, \quad j \in I, \quad k \in I_0,$$

$$z = (z_1, z_2, \cdots, z_n)^t, \quad z_j = \int_{\Gamma} g_2^\delta \varphi_j \, ds, \quad L = S + M.$$

In the following, we use the same symbol for a finite element function and its vector representation associated with the given finite element basis functions. Then, the finite element solutions $u_D^h$ and $u_N^h$ of the forward problems (93) and (94) corresponding to the source $f^k$, can be calculated by

$$L u_D^k = M_0 f^k, \quad u_{D,i}^k = g_1^\delta(x_{i}), \quad i \in I_b, \quad u_{D,i}^k = \sum_{i=1}^{n} u_{D,i} \varphi_i,$$

$$L u_N^k = M_0 f^k + z, \quad u_{N,i}^k = \sum_{i=1}^{n} u_{N,i} \varphi_i.$$

Similarly, for the discretization of the quality $K^*(K f - y^\delta)$, define

$$C = (c_{i,j}), \quad c_{i,j} = \int_{\Omega} \varphi_i \varphi_j \, dx, \quad i, j \in I.$$

Then the finite element approximation of $K^*(K f^k - y^\delta) = w_D^k - w_N^k$ can be calculated through

$$L w_D^k = C (u_D^k - u_N^k), \quad w_{D,i}^k = 0, \quad i \in I_b, \quad w_{D,i}^k = \sum_{i=1}^{n} w_{D,i} \varphi_i,$$

$$L w_N^k = C (u_D^k - u_N^k), \quad w_{N,i}^k = \sum_{i=1}^{n} w_{N,i} \varphi_i. \quad (95)$$

Note that (95) and (96) are the finite element discretization of the adjoint problems (68) and (69) with $u_{DN}$ being replaced by $u_D^k - u_N^k$. 

**Accelerated regularization methods**

25
References

[1] M. Abramowitz and I. Stegun. *Handbook of Mathematical Functions*. New York: Dover, 1972.

[2] H. Attouch, Z. Chbani, J. Peypouquet, and P. Redont. Fast convergence of inertial dynamics and algorithms with asymptotic vanishing viscosity. *Math. Program.*, 168:123–175, 2018.

[3] R. Bot, G. Dong, P. Elbau, and O. Scherzer. Convergence rates of first and higher order dynamics for solving linear ill-posed problems. *arXiv*, page 1812.09343, 2018.

[4] X. Cheng, R. Gong, W. Han, and W. Zheng. A novel coupled complex boundary method for inverse source problems. *Inverse Problems*, 30:055002, 2014.

[5] W. Du, Y. Wang, Q. Luo, and B. Liu. Optical molecular imaging for systems biology: from molecule to organism. *Anal. Bioanal. Chem.*, 386:444–457, 2006.

[6] H. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Dordrecht: Kluwer Academic Publishers Group, 1996.

[7] L. Evans. *Partial Differential Equations*. Providence: American Mathematical Society, 1998.

[8] R. Gong, X. Cheng, and W. Han. Theoretical analysis and numerical realization of bioluminescence tomography. *J. Concrete Appl. Math.*, 8:504–527, 2010.

[9] R. Gong, X. Cheng, and W. Han. A fast solver for an inverse problem arising in bioluminescence tomography. *J. Comp. Appl. Math.*, 267:228–243, 2014.

[10] R. Gong, X. Cheng, and W. Han. A new coupled complex boundary method for bioluminescence tomography. *Commun. Comput. Phys.*, 19:225–250, 2016.

[11] E. Hairer, G. Wanner, and C. Lubich. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations (Second Edition)*. New York: Springer, 2006.

[12] W. Han, W. Cong, and G. Wang. Mathematical theory and numerical analysis of bioluminescence tomography. *Inverse Problems*, 22:1659–1675, 2006.

[13] W. Han, J. Eichholz, J. Huang, and J. Lu. RTE-based bioluminescence tomography: A theoretical study. *Inv. Probl. Sci. Eng.*, 19:435–459, 2011.

[14] M. Hochbruck, C. Lubich, and H. Selhofer. Exponential integrators for large systems of differential equations. *SIAM J. Sci. Comput.*, 19:1152–1174, 1998.

[15] P. Mathé and S. V. Pereverzev. Geometry of linear ill-posed problems in variable Hilbert scales. *Inverse Problems*, 19(3):789–803, 2003.

[16] M. Nashed. A new approach to classification and regularization of ill-posed operator equations. In Inverse and Ill-posed Problems (Sankt Wolfgang, 1986), volume 4 of Notes Rep. Math. Sci. Engrgr., pages 53–75. Academic Press, Boston, MA, 1987.

[17] F. Natterer and F. Wübbeling. *Mathematical Methods in Image Reconstruction*. Philadelphia, PA: SIAM, 2001.

[18] A. Neubauer. On Nesterov acceleration for Landweber iteration of linear ill-posed problems. *J. Inverse Ill-Posed Probl.*, 25:381–390, 2017.

[19] V. Ntziachristos, C. Tung, C. Bremer, and R. Weissleder. Fluorescence molecular tomography resolves protease activity in vivo. *Nat. Med.*, 8:757–761, 2002.

[20] B. Rice, M. Cable, and M. Neison. In vivo imaging of light-emitting probes. *J. Biomed Opt.*, 6:432–440, 2001.

[21] A. Rieder. Runge-Kutta integrators yield optimal regularization schemes. *Inverse Problems*, 21:453–471, 2005.

[22] S. Song and J. Huang. Solving an inverse problem from bioluminescence tomography by minimizing an energy-like functional. *J. Comput. Anal. and Appl.*, 14:544–558, 2012.

[23] W. Su, S. Boyd, and E. Candès. A differential equation for modeling Nesterov’s accelerated gradient method: theory and insights. *J. Mach. Learn. Res.*, 17:1–43, 2016.

[24] U. Tautenhahn. On the asymptotical regularization of nonlinear ill-posed problems. *Inverse Problems*, 10:1405–1418, 1994.

[25] G. Vainikko and A. Veretennikov. *Iteration Procedures in Ill-Posed Problems*. Moscow: Nauka (In Russian), 1986.

[26] Y. Zhang, R. Gong, X. Cheng, and M. Gulliksson. A dynamical regularization algorithm for solving inverse source problems of elliptic partial differential equations. *Inverse Problems*, 34:065001, 2018.

[27] Y. Zhang, R. Gong, M. Gulliksson, and X. Cheng. A coupled complex boundary expanding compacts method for inverse source problems. *J. Inverse Ill-Posed Probl.*, 27:67–86, 2019.
[28] Y. Zhang and B. Hofmann. On fractional asymptotical regularization of linear ill-posed problems in Hilbert spaces. arXiv, page 1812.09734, 2018.

[29] Y. Zhang and B. Hofmann. On the second order asymptotical regularization of linear ill-posed inverse problems. Appl. Anal., DOI:10.1080/00036811.2018.1517412, 2018.