



Second order asymptotical regularization methods for inverse problems in partial differential equations

Ye Zhang^{a,b}, Rongfang Gong^{c,*}

^a Shenzhen MSU-BIT University, 518172 Shenzhen, China

^b School of Mathematics and Statistics, Beijing Institute of Technology, 100081 Beijing, China

^c Department of Mathematics, Nanjing University of Aeronautics and Astronautics, 211106 Nanjing, China



ARTICLE INFO

Article history:

Received 23 January 2019

Received in revised form 1 October 2019

Keywords:

Inverse source problems

Partial differential equations

Asymptotical regularization

Convergence

Finite element methods

Symplectic methods

ABSTRACT

We develop Second Order Asymptotical Regularization (SOAR) methods for solving inverse source problems in elliptic partial differential equations with both Dirichlet and Neumann boundary data. We show the convergence results of SOAR with the fixed damping parameter, as well as with a dynamic damping parameter, which is a continuous analog of Nesterov's acceleration method. Moreover, by using Morozov's discrepancy principle together with a newly developed total energy discrepancy principle, we prove that the approximate solution of SOAR weakly converges to an exact source function as the measurement noise goes to zero. A damped symplectic scheme, combined with the finite element method, is developed for the numerical implementation of SOAR, which yields a novel iterative regularization scheme for solving inverse source problems. Several numerical examples are given to show the accuracy and the acceleration effect of SOAR. A comparison with the state-of-the-art methods is also provided.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

In this paper, inspired by the asymptotical regularization [1–3], we establish a new framework for stably solving inverse problems in partial differential equations (PDEs). To present the ideas, we take the following inverse source problem as an example: given g_1 and g_2 on Γ , find p such that (p, u) satisfies

$$\begin{cases} -\Delta u + u = p\chi_{\Omega_0} & \text{in } \Omega, \\ u = g_1 \text{ and } \frac{\partial u}{\partial \mathbf{n}} = g_2 & \text{on } \Gamma, \end{cases} \quad (1)$$

where $\Omega \subset \mathbb{R}^d$ ($d = 2, 3$) represents a bounded domain with a smooth boundary Γ , $\partial/\partial \mathbf{n}$ stands for the unit outward normal derivative, $\Omega_0 \subset \Omega$ is known as a permissible region of the source function, and χ is the indicator function such that $\chi_{\Omega_0}(x) = 1$ for $x \in \Omega_0$, while $\chi_{\Omega_0}(x) = 0$, when $x \notin \Omega_0$. Over the last few decades, inverse source problems in PDEs have attracted a great attention, and numerous methods have been developed, including both theoretical approaches and numerical algorithms. We refer to the monograph [4] for more theoretical introductions. Some recently developed theoretical and numerical methods on solving the inverse source problems can be found in [5–8], and references therein. Note that the framework proposed in this paper can also be applied to various linear and nonlinear inverse problems in PDEs, e.g. inverse source problems in parabolic or hyperbolic PDEs, parameter identification problems in PDEs, etc.

* Corresponding author.

E-mail addresses: ye.zhang@smbu.edu.cn (Y. Zhang), grf_math@nuaa.edu.cn (R. Gong).

The variational methods of solving (1) are usually classified into two groups: the boundary fitting formulation and the domain fitting formulation. For the boundary fitting formulation, we use one of the boundary conditions to form a boundary value problem, and the remaining boundary condition as the object-optimized function to determine the source term. For instance, the following formulation can be considered [9]

$$\min_p \frac{1}{2} \|u(p) - g_1\|_{0,\Gamma}^2, \quad (2)$$

where $u(p)$ is the weak solution in $H^1(\Omega)$ of (1) with the Neumann boundary condition, and $\|\cdot\|_{0,\Gamma}$ is the standard norm of $L^2(\Gamma)$.

The Kohn–Vogelius method is certainly the most prominent domain fitting formulation for the inverse source problem (1). In this approach, the following optimization problem is adopted [10,11]:

$$\min_p \frac{1}{2} \|u_1(p) - u_2(p)\|_{0,\Omega}^2, \quad (3)$$

where $u_1, u_2 \in H^1(\Omega)$ are the weak solutions of $-\Delta u_{1,2} + u_{1,2} = p\chi_{\Omega_0}$ with Dirichlet and Neumann data respectively, and $\|\cdot\|_{0,\Omega}$ is the standard norm of $L^2(\Omega)$.

However, both formulations (2) and (3) use the Neumann and Dirichlet data separately. In [12], a novel coupled complex boundary method (CCBM) was introduced. The idea of CCBM is to couple the Neumann data and Dirichlet data in a Robin boundary condition, which leads to the following optimization problem

$$\min_p \frac{1}{2} \|u_{im}\|_{0,\Omega}^2. \quad (4)$$

where $u = u_{re} + iu_{im}$ ($i = \sqrt{-1}$ is the imaginary unit) solves (in the weak form)

$$\begin{cases} -\Delta u + u = p\chi_{\Omega_0} & \text{in } \Omega, \\ \frac{\partial u}{\partial \mathbf{n}} + iu = g_2 + ig_1 & \text{on } \Gamma. \end{cases} \quad (5)$$

Obviously, all formulations (2), (3) and (4) are still ill-posed, since a general source could not be determined uniquely by a single boundary measurement, see e.g., [4,13]. Moreover, the mapping from the source function to the boundary data is a compact operator in Hilbert spaces, which implies the unboundedness of its inversion operator. Therefore, for the problem with noisy boundary data, regularization methods should be employed for obtaining stable approximate solutions. Loosely speaking, three groups of regularization methods exist: descriptive regularization methods, variational regularization methods and iterative regularization methods.

Descriptive regularization uses a priori information of the solution to overcome the ill-posedness of the original inverse problem. For inverse source problems, under the assumption of sourcewise representation of the unknown source function, the authors in [14] combined the expanding compacts method and CCBM to propose a new efficient regularization method. However, in this paper, we are interested in a more general case that no a priori information about the solution is available.

Tikhonov regularization should be the most prominent variational regularization method. Denote $V(p)$ as the objective functional in (2), (3) or (4). With the Tikhonov regularization, the original inverse source problem (1) is converted to the following minimization problem:

$$p_\varepsilon = \arg \min_p V_\varepsilon(p), \quad V_\varepsilon(p) := V(p) + \frac{\varepsilon}{2} \|p\|_{0,\Omega_0}^2, \quad (6)$$

where $\varepsilon > 0$ is a regularization parameter chosen in a special way using the noisy boundary data. Under certain assumptions, (6) admits a unique solution p_ε , which converges to the minimal norm solution of (1) with the noise-free boundary data [9,10,12].

In this paper, our focus is on the iterative regularization approaches, since, from a computational viewpoint, the iterative approach seems more tractable, especially for large-scale problems. The most famous iterative regularization approach should be the Landweber iteration, which is defined by (cf., e.g., [15,16])

$$p_{k+1} = p_k - \Delta t \nabla V(p_k), \quad (7)$$

which can be viewed as a discrete analog of the following first order evolution equation

$$\dot{p}(t) = -\nabla V(p(t)), \quad (8)$$

where ∇ denotes the gradient of V , and t is the introduced artificial time. The formulation (8) is known as the asymptotical regularization, or the Showalter's method. The regularization property of (8) can be analyzed through a proper choice of the terminating time.

It is well known that the Landweber method works quite slowly. Thus, accelerating strategies are usually adopted in practice. In recent years, there has been increasing evidence to show that the second order iterative methods exhibit remarkable acceleration properties for stably solving ill-posed problems. The most well-known methods are the Nesterov acceleration scheme [17], the ν -method [15, § 6.3], and the two-point gradient method [18]. Recently, the authors

in [3,19] have established an initial theory of the second order asymptotical regularization method for solving general linear ill-posed inverse problems. In this paper, inspired by the development of second order dynamics for accelerating the convergence of iterative regularization methods in [3,18,19], we develop a second order asymptotical regularization method for solving the inverse source problem (1), i.e., we consider the second order evolution equation

$$\begin{cases} \ddot{p}(t) + \eta(t)\dot{p}(t) + \nabla V(p(t)) = 0, \\ p(0) = p_0, \quad \dot{p}(0) = \dot{p}_0, \end{cases} \quad (9)$$

where $(p_0, \dot{p}_0) \in P \times P$ is the prescribed initial data, $\eta > 0$ is the so-called damping parameter, which may or may not depend on the artificial time t , and P is the solution space, which will be precisely defined later. It is not difficult to show that the evolution equation (9) with the following specific choice of discretization parameters

$$\begin{cases} \Delta t_k = 4 \frac{(2k+2v-1)(k+v-1)}{(k+2v-1)(2k+4v-1)}, \\ \eta_k = \frac{(k+2v-1)(2k+4v-1)(2k+2v-3) - (k-1)(2k-3)(3k+3v-1)}{4(2k+2v-3)(2k+2v-1)(k+v-1)}, \end{cases}$$

yields the v -method. Moreover, as demonstrated in [20], (9) with a special choice of damping parameter can be considered as an infinite dimensional extension of the Nesterov's scheme in the following sense.

Theorem 1. Let $\{p_k\}$ be the sequence, generated by the Nesterov's scheme with parameters (α, ω) , see (70) for details. Then, for all fixed $T > 0$:

$$\lim_{\omega \rightarrow 0} \max_{0 \leq k \leq T/\sqrt{\omega}} \|p_k - p(k\sqrt{\omega})\|_P = 0,$$

where $p(\cdot)$ is the solution of (9) with $\eta(t) = \alpha/t$.

We highlight the main contributions of our work in the following:

- We develop a new continuous regularization method (9) for inverse problems in elliptic PDEs. Based on the numerical discretization of (9), a new class of iterative regularization methods for solving inverse source problems has been created. To the best of our knowledge, only a few of iterative regularization methods exist for solving inverse problems in PDEs. Therefore, the newly developed continuous regularization method, as well as the associated iterative regularization methods, will be a competitive candidate for studying various inverse problems in PDEs.
- Based on our new framework of regularization, we develop a novel a posteriori regularization parameter selection method – the total energy discrepancy principle. Due to the difficulty of numerical implementation, until recently the theoretical results of most of the works in PDE-controlled inverse problems are limited to the a priori regularization parameter choice rule. Our newly developed a posteriori regularization parameter choice method is based on a class of iterative regularization methods, and can be simply used during the iteration. Therefore, the numerical realization of our regularization parameter choice rule is simple and this makes it can be easily adopted in practice.
- As far as we know, only a few accelerated regularization methods exist in the literature of inverse problems. Our numerical experiments indicate that in comparison with conventional accelerated iterative regularization methods, the developed new regularization methods exhibit good acceleration effect for the presented PDE-controlled inverse problem. Therefore, the proposed new class of iterative regularization methods significantly extends the class of fast regularization methods for solving large scale inverse problems in PDEs.

The remainder of the paper is structured as follows: Section 2 discusses some properties of the solution of evolution equation (9). The convergence analysis for exact and noisy data are presented in Sections 3 and 4, respectively. Finite dimensional approximation of our method is proposed in Section 5, where we develop a novel second order iterative regularization algorithm. Some numerical examples, as well as a comparison with three existing iterative regularization methods, are presented in Section 6. Finally, concluding remarks are given in Section 7.

2. Properties of the second order evolution equation

For clarity, we only consider the formulation (4) in this paper. Let us first introduce the notations for the function spaces that are used in this paper. For a set G (e.g., Ω , Ω_0 or Γ), denote by $W^{m,s}(G)$ the Sobolev space with norm $\|\cdot\|_{m,s,G}$ [21]. In particular, $L^s(G) := W^{0,s}(G)$. Moreover, $H^m(G)$ represents $W^{m,2}(G)$ with the corresponding inner product $(\cdot, \cdot)_{m,G}$ and norm $\|\cdot\|_{m,G}$. Let $\mathbf{H}^m(G)$ be the complex version of $H^m(G)$ with inner product $((\cdot, \cdot))_{m,G}$ and norm $\|\cdot\|_{m,G}$ defined as follows: $\forall u, v \in \mathbf{H}^m(G)$, $((u, v))_{m,G} = (u, \bar{v})_{m,G}$, $\|u\|_{m,G} = ((u, u))_{m,G}^{1/2}$, where \bar{v} is the conjugate complex of v . Denote $P = L^2(\Omega_0)$ or $H^1(\Omega_0)$ as the space for the source function p . Its corresponding inner product and norm are given by $(\cdot, \cdot)_P$ and $\|\cdot\|_P$, respectively.

Assume that $g_1 \in H^{1/2}(\Gamma) \cap L^\infty(\Gamma)$ and $g_2 \in L^\infty(\Gamma)$. Moreover, instead of the exact data $\{g_1, g_2\}$, we have only the noisy data $g_1^\delta, g_2^\delta \in L^\infty(\Gamma)$ such that

$$\|g_1^\delta - g_1\|_{\infty,\Gamma} \leq \delta, \quad \|g_2^\delta - g_2\|_{\infty,\Gamma} \leq \delta, \quad (10)$$

where $\delta > 0$ denotes the error level of the measurement. Then, the CCBM for inverse source problem (1) with noisy data $\{g_1^\delta, g_2^\delta\}$ can be formulated as

$$\inf_{p \in P} V(p) = \inf_{p \in P} V(p; \delta) = \inf_{p \in P} \frac{1}{2} \|u_{im}(p)\|_{0,\Omega}^2, \quad (11)$$

where $u = u_{re} + iu_{im}$ solves

$$\begin{cases} -\Delta u + u = p\chi_{\Omega_0} & \text{in } \Omega, \\ \frac{\partial u}{\partial \mathbf{n}} + iu = g_2^\delta + ig_1^\delta & \text{on } \Gamma. \end{cases} \quad (12)$$

Suppose that system (1) has at least one solution for noise-free data and denote by p^\dagger one of the solutions. Obviously, for all $p \in P$: $V(p; 0) \geq 0 = \frac{1}{2} \|u_{im}(p^\dagger)\|_{0,\Omega}^2 = V(p^\dagger; 0)$. Hence, we have

$$p^\dagger \in \arg \min_{p \in P} V(p; 0). \quad (13)$$

Proposition 1 ([22, Proposition 1]). The Fréchet derivative of $V(p)$, defined in (11), is the imaginary part of the solution to the adjoint problem

$$\begin{cases} -\Delta w + w = u_{im}(p) & \text{in } \Omega, \\ \frac{\partial w}{\partial \mathbf{n}} + iw = 0 & \text{on } \Gamma, \end{cases} \quad (14)$$

where u_{im} is the imaginary part of u , the solution of (12), i.e., $\nabla_p V(p) = w_{im}(p)\chi_{\Omega_0}$.

It is not difficult to show that $V''(p)q^2 = \|u_{im}(q) - u_{im}(0)\|_{0,\Omega}^2$. Hence, $V(p)$ is convex.

Now we are in a position to introduce the second order asymptotical regularization for solving the inverse source problem (1).

Definition 1. Let $p^\delta(\cdot, t) \in P$ be the solution to the following Cauchy problem

$$\begin{cases} \ddot{p}^\delta(x, t) + \eta(t)\dot{p}^\delta(x, t) + w_{im}(x, t) = 0, & x \in \Omega_0, t \in (0, \infty), \\ p^\delta(x, 0) = p_0(x), \dot{p}^\delta(x, 0) = \dot{p}_0(x), & x \in \Omega_0, \end{cases} \quad (15)$$

where $w = w_{re} + iw_{im}$ is the solution of the adjoint problem with the same t

$$\begin{cases} -\Delta w(x, t) + w(x, t) = u_{im}(p^\delta(x, t)), & x \in \Omega, t \in (0, \infty), \\ \frac{\partial w(x, t)}{\partial \mathbf{n}} + iw(x, t) = 0, & x \in \Gamma, t \in (0, \infty), \end{cases} \quad (16)$$

and $u = u_{re} + iu_{im}$ is the solution of the BVP

$$\begin{cases} -\Delta u(x, t) + u(x, t) = p^\delta(x, t)\chi_{\Omega_0}, & x \in \Omega, t \in (0, \infty), \\ \frac{\partial u(x, t)}{\partial \mathbf{n}} + iu(x, t) = g_2^\delta(x) + ig_1^\delta(x), & x \in \Gamma, t \in (0, \infty). \end{cases} \quad (17)$$

Then, $p^\delta(x, T^*)$, equipped with an appropriate terminating time point $T^* = T^*(\delta)$, is called a second order asymptotical regularized solution of (1) if $p^\delta(\cdot, T^*)$ (strong or weakly) converges to p^\dagger in P as $\delta \rightarrow 0$.

Before presenting the solvability of system (15)–(17), we discuss the well-posedness of the BVPs (16) and (17). For any $u, \psi \in \mathbf{H}^1(\Omega)$, define

$$\begin{aligned} a(u, \psi) &= \int_{\Omega} (\nabla u \cdot \nabla \bar{\psi} + u\bar{\psi}) dx + i \int_{\Gamma} u\bar{\psi} ds, \\ f^\delta(\psi) &= \int_{\Omega_0} p^\delta \bar{\psi} dx + \int_{\Gamma} g_2^\delta \bar{\psi} ds + i \int_{\Gamma} g_1^\delta \bar{\psi} ds. \end{aligned}$$

Then the weak form of the BVP (17) reads:

$$\text{find } u \in \mathbf{H}^1(\Omega) \text{ such that } a(u, \psi) = f^\delta(\psi), \quad \forall \psi \in \mathbf{H}^1(\Omega). \quad (18)$$

Lemma 2 ([12]). Problem (18) admits a unique solution $u \in \mathbf{H}^1(\Omega)$ which depends continuously on p^δ, g_1^δ and g_2^δ . Furthermore, a constant $C(\Omega)$ exists such that

$$\|u\|_{1,\Omega} \leq C(\Omega) (\|p^\delta\|_{0,\Omega_0} + \|g_1^\delta\|_{0,\Gamma} + \|g_2^\delta\|_{0,\Gamma}). \quad (19)$$

By Lemma 2 and the definition of $V(p)$ in (11) and w in (14), it is not difficult to prove the following lemma.

Lemma 3. The following two inequalities hold for some constants $C(\Omega)$:

$$V(p^\delta) \leq C(\Omega) (\|p^\delta\|_{0,\Omega_0}^2 + \|g_1^\delta\|_{0,\Gamma}^2 + \|g_2^\delta\|_{0,\Gamma}^2), \quad (20)$$

$$\|w(p^\delta)\|_{1,\Omega} \leq C(\Omega) (\|p^\delta\|_{0,\Omega_0} + \|g_1^\delta\|_{0,\Gamma} + \|g_2^\delta\|_{0,\Gamma}). \quad (21)$$

Theorem 4. For each pair $(p_0, \dot{p}_0) \in P \times P$, system (15)–(17) has a unique solution $p^\delta(\cdot, t)$ which depends continuously on the boundary data $\{g_1^\delta, g_2^\delta\}$.

The proof is similar to those of (a) in [22, Theorem 1]. A sketch of the proof is given in Appendix A.

3. Convergence for noise-free boundary data

In this section, we investigate two models: when the damping parameter η is fixed, and when it is time dependent. For simplicity, let $p(t) = p(\cdot, t)$.

3.1. Case I: η is a constant

We first study the dynamics of the solution $p(t) \in P$ of system (15)–(17).

Lemma 5. Let $p(t)$ be the solution of (15)–(17) with the exact data $\{g_1, g_2\}$. Then, in the case $\eta \geq 1$, we have

- (i) $p \in L^\infty([0, \infty), P)$.
- (ii) $\dot{p} \in L^\infty([0, \infty), P) \cap L^2([0, \infty), P)$ and $\dot{p}(\cdot, t) \rightarrow 0$ as $t \rightarrow \infty$.
- (iii) $\ddot{p} \in L^\infty([0, \infty), P) \cap L^2([0, \infty), P)$ and $\ddot{p}(\cdot, t) \rightarrow 0$ as $t \rightarrow \infty$.
- (iv) $V(p(t)) = o(t^{-1})$ as $t \rightarrow \infty$.

The proof of above lemma follows the idea in [23], and can be found in Appendix B. The rate $V(p(t)) = o(t^{-1})$ as $t \rightarrow \infty$ given in Lemma 5 for the second order evolution equation (9) should be compared with the corresponding result for the first order method, i.e. the gradient decent methods, where one only obtains $V(p(\cdot, t)) = \mathcal{O}(t^{-1})$ as $t \rightarrow \infty$. If we consider a discrete iterative method with the number k of iterations, assertion (iv) in Lemma 5 indicates that in comparison with gradient descent methods, the second order methods (9) need the same computational complexity for the number k of iterations, but can achieve a higher order $o(k^{-1})$ of accuracy of the objective functional.

Now, we list the following two lemmas, which will be used in the convergence analysis of the dynamical solution $p(x, t)$.

Lemma 6 (Opial Lemma [23, Lemma 4.1], [24]). Let P be a Hilbert space and $p : [0, \infty) \rightarrow P$ be a mapping such that there exists a non-empty set $S \subset P$ which satisfies

- (i) $\forall t_n \rightarrow \infty$ with $p(t_n) \rightharpoonup \bar{p}$ weakly in P , we have $\bar{p} \in S$.
- (ii) $\forall p^\dagger \in S$, $\lim_{t \rightarrow \infty} \|p(t) - p^\dagger\|_P$ exists.

Then, $p(t)$ weakly converges as $t \rightarrow \infty$ to some element of S .

Lemma 7 (Lemma 4.2 in [23]). Let $\varphi(t) \in C^1((0, \infty), [0, +\infty))$ satisfy the inequality $\ddot{\varphi}(t) + \eta\dot{\varphi}(t) \leq g(t)$ with $g(t) \in L^1((0, \infty), [0, +\infty))$. Then, $\dot{\varphi}_+$, the positive part of $\dot{\varphi}$, belongs to $L^1((0, \infty), [0, +\infty))$ and, as a consequence, $\lim_{t \rightarrow \infty} \varphi(t)$ exists.

Now, we are in a position to present the main result in this section.

Theorem 8. The solution $p(t)$ of (15)–(17) with the exact data converges weakly in P to an exact source function of inverse source problem (1) as $t \rightarrow \infty$.

Proof. It suffices to check two conditions in Opial lemma. Denote by $S := \arg \min_{p \in P} V(p; 0)$. It follows from (13) that $S \neq \emptyset$. Consider a sequence $\{p(t_n)\}$ such that $p(t_n) \rightharpoonup \bar{p}$ weakly in P . Applying the convexity inequality to the functional $V(p) = \frac{1}{2} \|u_{im}\|_{0,\Omega}^2$ we have

$$V(z) \geq V(p(t_n)) + (z - p(t_n), \nabla V(p(t_n)))_P, \quad \forall z \in P. \quad (22)$$

By using the continuity of $V(p)$, and noticing that, in the inner product $(z - p(t_n), \nabla V(p(t_n)))_P$, the two terms are, respectively, norm converging to zero and weakly convergent, we can pass to the lower limit to obtain $V(z) \geq V(\bar{p})$ for all $z \in P$. Set $z = p^\dagger$ in the above inequality, we conclude that $0 = V(p^\dagger) \geq V(\bar{p})$, which implies that $\bar{p} \in S$.

Now, we prove the second requirement in Lemma 6. It is equivalent to show that $\lim_{t \rightarrow \infty} e(t)$ exists, where $e(t) = \frac{1}{2} \|p(t) - p^\dagger\|_P^2$ is defined in the proof of Lemma 5. From (B.3), we deduce that

$$\ddot{e}(t) + \eta\dot{e}(t) \leq \|\dot{p}(t)\|_P^2. \quad (23)$$

Since $\dot{p}(\cdot) \in L^2([0, \infty), P)$, inequality (23) together with Lemma 7 yields the second condition in Opial lemma. This completes the proof of the weak convergence of the dynamical solution of (9). \square

3.2. Case II: $\eta(t) = r/t$

Now, we study the second order dynamical system (15)–(17) with an asymptotical vanishing damping parameter of the type $\eta(t) = r/t$, i.e. we consider the following evolution equation

$$\begin{cases} \ddot{p}(x, t) + \frac{r}{t}\dot{p}(x, t) + w_{im}(x, t) = 0, & x \in \Omega_0, \quad t \in (1, \infty), \\ p(x, 1) = p_0(x), \quad \dot{p}(x, 1) = \dot{p}_0(x), & x \in \Omega_0, \end{cases} \quad (24)$$

where $w = w_{re} + iw_{im}$ is the solution of the adjoint problem (16) with the same t . As discussed in Section 1, this is a particularly interesting case as the second order flow (24) yields a continuous version of Nesterov's scheme, which has a higher order of convergence rate for the residual functional, i.e. $V(p(t)) = \mathcal{O}(k^{-2})$ for $r = 3$ and $V(p(t)) = o(k^{-2})$ for $r > 3$ [25].

Remark 1. We shift the initial time point from 0 to 1 for the regularity of the term r/t . Otherwise, one can use $r/(t+1)$ instead of r/t in (24).

For our further use, we introduce the anchored energy function

$$\varepsilon_\lambda(t) = t^2 V(p(t)) + \frac{1}{2} \|\lambda(p(t) - p^\dagger) + t\dot{p}(t)\|_P^2 + \frac{\lambda(r-1-\lambda)}{2} \|p(t) - p^\dagger\|_P^2, \quad (25)$$

where the exact source p^\dagger is given in (13). For $r \geq 3$, using the convexity inequality $0 = V(p^\dagger) \geq V(p) + (\nabla V(p), p^\dagger - p)_P$ for all $p \in P$ and (24), it is not difficult to show that

$$\dot{\varepsilon}_\lambda(t) \leq -(\lambda-2)tV(p(t)) - (r-1-\lambda)t\|\dot{p}(t)\|_P^2. \quad (26)$$

Hence, for $r \geq 3$ and $\lambda \in [2, r-1]$, $\varepsilon_\lambda(t)$ is non-increasing.

Now, we are in a position to derive similar results to those in Section 3.1.

Lemma 9. Let $p(t)$ be the solution of (24) with the exact data and parameter $r > 3$. Then, $\dot{p} \in L^\infty([1, \infty), P) \cap L^2([1, \infty), P)$ and $\dot{p}(t) \rightarrow 0$ as $t \rightarrow \infty$. Moreover, $V(p(t)) = \mathcal{O}(t^{-2})$ as $t \rightarrow \infty$.

Proof. This proof uses the technique in [26]. Consider the Lyapunov function of (24) by $\varepsilon(t) = \frac{1}{2} \|\dot{p}(t)\|_P^2 + V(p(t))$. It is easy to show that

$$\dot{\varepsilon}(t) = -\frac{r}{t} \|\dot{p}(t)\|_P^2 \leq 0. \quad (27)$$

Hence, $\varepsilon(t)$ is non-increasing, and $\varepsilon(\infty) := \lim_{t \rightarrow \infty} \varepsilon(t)$ exists by noting that $\varepsilon(t) \geq 0$ for all t . Furthermore, by $\|\dot{p}(t)\|_P^2 \leq 2\varepsilon(t) \leq 2\varepsilon(1)$ we conclude the uniform boundedness of $\dot{p}(\cdot)$.

Set $\lambda = 2$ in (26) to derive

$$\dot{\varepsilon}_2(t) \leq -(r-3)t\|\dot{p}(t)\|_P^2. \quad (28)$$

Integrating both sides in (28), we obtain that

$$\int_1^\infty \|\dot{p}(t)\|_P^2 dt \leq \int_1^\infty t\|\dot{p}(t)\|_P^2 dt \leq \varepsilon_2(1)/(r-3) < \infty, \quad (29)$$

which yields $\dot{p}(\cdot) \in L^2([1, \infty), P)$. Now, consider the function $e(t) = \frac{1}{2} \|p(t) - p^\dagger\|_P^2$. Using the local convexity of $V(\cdot)$ and Eq. (9), similar to (B.3), it is not difficult to obtain

$$\ddot{e}(t) + \frac{r}{t}\dot{e}(t) + V(p(t)) \leq \|\dot{p}(t)\|_P^2. \quad (30)$$

Divide this expression by t to obtain

$$\frac{1}{t}\ddot{e}(t) + \frac{r}{t^2}\dot{e}(t) + \frac{1}{t}V(p(t)) \leq \frac{3}{2t}\|\dot{p}(t)\|_P^2.$$

Integrating above inequality from 1 to t and using integration by parts for $\ddot{e}(t)$, we obtain

$$\int_1^t \frac{\varepsilon(\tau)}{\tau} d\tau \leq \dot{e}(1) - \frac{\dot{e}(t)}{t} - (r+1) \int_1^t \frac{\dot{e}(\tau)}{\tau^2} d\tau + \frac{3}{2} \int_1^t \frac{\|\dot{p}(\tau)\|_P^2}{\tau} d\tau. \quad (31)$$

On one hand, using the integration by parts and the positivity of functional $e(\cdot)$, we have

$$\int_1^t \frac{\dot{e}(\tau)}{\tau^2} d\tau = \frac{e(t)}{t^2} - e(1) + 2 \int_1^t \frac{e(\tau)}{\tau^3} d\tau \geq -e(1). \quad (32)$$

On the other hand, relation (27) gives

$$\int_1^t \frac{\|\dot{p}(\tau)\|_P^2}{\tau} d\tau = \frac{\varepsilon(1) - \varepsilon(t)}{r}. \quad (33)$$

Combine (31)–(33) to get

$$\int_1^t \frac{\mathcal{E}(\tau)}{\tau} d\tau \leq \dot{e}(1) - \frac{\dot{e}(t)}{t} + (r+1)e(1) + \frac{3(\mathcal{E}(1) - \mathcal{E}(t))}{2r} = C(1) - \frac{\dot{e}(t)}{t} - \frac{3\mathcal{E}(t)}{2r}, \quad (34)$$

where $C(1) = \dot{e}(1) + (r+1)e(1) + \frac{3\mathcal{E}(1)}{2r}$ collects all constant terms. For any $T \geq t > 1$, we have

$$\mathcal{E}(T) \int_1^t \frac{1}{\tau} d\tau + \frac{3\mathcal{E}(T)}{2r} \leq C(1) - \frac{\dot{e}(t)}{t} \quad (35)$$

by noting the non-increasing of Lyapunov function $\mathcal{E}(t)$. Rewrite (35) as $\mathcal{E}(T) (\ln(t) + \frac{3}{2r}) \leq C(1) - \frac{\dot{e}(t)}{t}$, and then integrate it from $t = 1$ to $t = T$ to have

$$\mathcal{E}(T) \left(T \ln(T) + 1 - T + \frac{3}{2r}(T-1) \right) \leq C(1)(T-1) - \int_1^T \frac{\dot{e}(t)}{t} dt. \quad (36)$$

Moreover, using the integration by parts and the positivity of functional $e(\cdot)$, we have

$$\int_1^T \frac{\dot{e}(t)}{t} dt = \frac{e(T)}{T} - e(1) + \int_1^T \frac{e(t)}{t^2} dt \geq -e(1). \quad (37)$$

By combining (36) and (37), we deduce that

$$\mathcal{E}(T) (T \ln(T) + C_1 T + C_2) \leq C(1)T + C_3, \quad (38)$$

where $C_1 = \frac{3}{2r} - 1$, $C_2 = 1 - 3/(2r)$ and $C_3 = e(1) - C(1)$ are three constants.

Inequality (38) immediately yields $\mathcal{E}(\infty) \leq 0$. By the non-negativity of Lyapunov function $\mathcal{E}(\cdot)$, we conclude $\mathcal{E}(\infty) = 0$, which implies that both $\dot{p}(T)$ and $V(p(T))$ converge to 0 in P when $T \rightarrow \infty$.

Finally, let us show the convergence rate of $V(p(t))$. Set $\lambda = r - 1$ in (25) to obtain $t^2 V(p(t)) \leq \mathcal{E}_{r-1}(t)$. Since $\mathcal{E}_{r-1}(t)$ is non-increasing, we conclude that $V(p(t)) \leq \mathcal{E}_{r-1}(1)/t^2$. \square

Lemma 10 (Lemma 5.9 in [26]). Let $\varphi(t) \in C^1((1, \infty), [0, +\infty))$ satisfy the inequality $t\ddot{\varphi}(t) + r\dot{\varphi}(t) \leq g(t)$ with $r \geq 1$ and $g(t) \in L^1((1, \infty), [0, +\infty))$. Then, $\dot{\varphi}_+$, the positive part of $\dot{\varphi}$, belongs to $L^1((1, \infty), [0, +\infty))$ and, as a consequence, $\lim_{t \rightarrow \infty} \varphi(t)$ exists.

Theorem 11. The solution $p(t)$ of (24) with $r > 3$ converges weakly to an exact source function of inverse source problem (1) in P as $t \rightarrow \infty$.

Proof. Set $\lambda = 2$ in (25) to derive $\|p(t) - p^\dagger\|_P^2 \leq \frac{\mathcal{E}_2(t)}{r-3} \leq \frac{\mathcal{E}_2(1)}{r-3}$, which yields the uniform boundedness of $p(t)$. Furthermore, from (29) we have $t\|\dot{p}(t)\|_P^2 \in L^1((1, \infty), [0, \infty))$. Recall from (30) to obtain $t\ddot{e}(t) + r\dot{e}(t) \leq t\|\dot{p}(t)\|_P^2$. From Lemma 10, and note that $t\|\dot{p}(t)\|_P^2$ is integrable on $[1, \infty)$, the limit $\lim_{t \rightarrow \infty} e(t)$ exists. This gives the second hypothesis in Opial's Lemma. The first one was established in Lemma 9, i.e. $V(p(t)) \rightarrow 0$ as $t \rightarrow \infty$. This completes the proof by using the Opial's Lemma 6. \square

Remark 2. (a) In Theorems 8 and 11, we only obtain the weak convergence for both fixed and dynamic damping parameters. One way to obtain the strong convergence result is to include a regularization term $\epsilon(t)p(x, t)$ in the evolution equation (9) with a specially chosen dynamic regularization parameter $\epsilon(t)$, see formula (68) and [22] for details. However, the numerical results in Section 6.4, cf. Table 5, show that our method works much better than this method in terms of accuracy and speed. (b) Let Π^h be any project operator, acting from P into a finite element space P^h for a fixed triangulation \mathcal{T}^h . Then, we have the strong convergence $\Pi^h p^\delta(t) \rightarrow \Pi^h p^\dagger(t)$ as $t \rightarrow \infty$ in P^h , since strong convergence and weak convergence coincide in any finite dimensional/element space. This fact will be used in Theorem 15 about the strong convergence of the finite element approximate solution.

4. Convergence for noisy data

In this section, we investigate the regularization property of the dynamic solution $p^\delta(t)$ of (15)–(17), equipped with some appropriate selection rules of the terminating time T^* .

Proposition 2. There exists a constant $C_0(\Omega)$, depending only on the geometry of the domain Ω , such that $\|u_{im}(p^\dagger)\|_{0,\Omega} \leq C_0(\Omega)\delta$, where $u = u_{re} + iu_{im}$ solves (12) and p^\dagger is defined in (13). Consequently, we have $V(p^\dagger) \leq C_0^2 \delta^2$. If Ω is a ball in \mathbb{R}^d centered at O with radius R or an annulus in \mathbb{R}^d centered at O with radius R and $r(< R)$, we have

$$C_0(\Omega) = \max(d, R)(2\pi)^{1/2}. \quad (39)$$

Proof. Denote by \tilde{u} the weak solution of (12) with the exact source term p^\dagger . Define $v := u - \tilde{u}$. Then v satisfies

$$\begin{cases} -\Delta v + v = 0 & \text{in } \Omega, \\ \frac{\partial v}{\partial \mathbf{n}} + iv = (g_2^\delta - g_2) + i(g_1^\delta - g_1) & \text{on } \Gamma. \end{cases} \quad (40)$$

The weak form of the above BVP (40) reads:

$$\text{find } v \in \mathbf{H}^1(\Omega) \text{ such that } a(v, \psi) = \tilde{f}^\delta(\psi), \quad \forall \psi \in \mathbf{H}^1(\Omega), \quad (41)$$

where $\tilde{f}^\delta(\psi) = \int_\Gamma (g_2^\delta - g_2) \bar{\psi} ds + i \int_\Gamma (g_1^\delta - g_1) \bar{\psi} ds$. Denote by v_{re} and v_{im} the real and imaginary parts of v , respectively. Obviously, $v_{im} \equiv u_{im}$ by noting $\tilde{u}_{im} = 0$. Furthermore, if one separates the real and imaginary parts of problem (40), the real part v_{re} of v satisfies

$$\begin{cases} -\Delta v_{re} + v_{re} = 0 & \text{in } \Omega, \\ \frac{\partial v_{re}}{\partial \mathbf{n}} - v_{im} = g_2^\delta - g_2 & \text{on } \Gamma, \end{cases}$$

whose weak form is

$$\int_\Omega (\nabla v_{re} \cdot \nabla \psi + v_{re} \psi) dx = \int_\Gamma (g_2^\delta - g_2) \psi ds + \int_\Gamma v_{im} \psi ds, \quad \forall \psi \in H^1(\Omega). \quad (42)$$

The imaginary part v_{im} of v satisfies

$$\begin{cases} -\Delta v_{im} + v_{im} = 0 & \text{in } \Omega, \\ \frac{\partial v_{im}}{\partial \mathbf{n}} + v_{re} = g_1^\delta - g_1 & \text{on } \Gamma, \end{cases}$$

whose weak form is

$$\int_\Omega (\nabla v_{im} \cdot \nabla \psi + v_{im} \psi) dx = \int_\Gamma (g_1^\delta - g_1) \psi ds - \int_\Gamma v_{re} \psi ds, \quad \forall \psi \in H^1(\Omega). \quad (43)$$

Set $\psi = v_{re}$ in (42) and $\psi = v_{im}$ in (43), and then add these two equations together to obtain

$$\|v_{re}\|_{1,\Omega}^2 + \|v_{im}\|_{1,\Omega}^2 = \int_\Gamma (g_2^\delta - g_2) v_{re} ds + \int_\Gamma (g_1^\delta - g_1) v_{im} ds,$$

which implies

$$\|v\|_{1,\Omega}^2 \leq \delta \int_\Gamma (|v_{re}| + |v_{im}|) ds. \quad (44)$$

On the other hand, if Ω is a ball/annulus in \mathbb{R}^d centered at 0 with radius R (and r), it holds [27]

$$\int_\Gamma |u(s)| ds \leq \frac{d}{R} \int_\Omega |u(x)| dx + \int_\Omega |\nabla u(x)| dx \quad (45)$$

for all $u \in W^{1,1}(\Omega)$. Then, by inequality (45) and the Cauchy-Schwarz inequality $\int_\Omega |u(x)| dx \leq R\pi^{1/2} \|u\|_{0,\Omega}$, we deduce that for $k = re$ or im

$$\int_\Gamma |v_k| ds \leq d\pi^{1/2} \|v_k\|_{0,\Omega} + R\pi^{1/2} \|\nabla v_k\|_{0,\Omega} \leq \max(d, R)\pi^{1/2} \|v_k\|_{1,\Omega}. \quad (46)$$

Combine (44), (46), and the inequality $\|v_{re}\|_{1,\Omega} + \|v_{im}\|_{1,\Omega} \leq \sqrt{2} \|v\|_{1,\Omega}$ to obtain

$$\|u_{im}(p^\dagger)\|_{0,\Omega} = \|v_{im}\|_{0,\Omega} \leq \|v\|_{1,\Omega} \leq \max(d, R)(2\pi)^{1/2} \delta,$$

which yields the required result. For the general smooth bounded domain, the proposition can be proven by using the Sobolev trace embedding inequality (with the constant S)

$$S \int_\Gamma |u(s)| ds \leq \int_\Omega |u(x)| + |\nabla u(x)| dx. \quad \square \quad (47)$$

Remark 3. The best (largest) embedding constant in (47) equals

$$S = \inf_{u \in W^{1,1}(\Omega) \setminus W_0^{1,1}(\Omega)} \frac{\int_\Omega |u(x)| + |\nabla u(x)| dx}{\int_\Gamma |u(s)| ds}. \quad (48)$$

The extrema of (48) exist as the embedding (47) is compact, cf. [28]. To the best of our knowledge, the rigorous lower bounds of S , hence the value of $C_0(\Omega)$ in Proposition 2, for general smooth domain Ω is still open. Alternatively, one can

estimate the value of S by numerically solving the following non-linear eigenvalue problem

$$\begin{cases} \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right) = 1 & \text{in } \Omega, \\ \frac{\partial u}{\partial \mathbf{n}} = \lambda |\nabla u| & \text{on } \Gamma, \end{cases} \quad (49)$$

by noting that the extrema in (48) can be assumed positive, see e.g., [29,30].

Proposition 3. Let $p^\delta(t)$ be the dynamic solution of (15)–(17) with the fixed damping parameter $\eta \geq 1$ or $\eta(t) = r/t$ ($r > 3$). Then, $\lim_{t \rightarrow \infty} V(p^\delta(t)) \leq C_0^2 \delta^2$, where $V(\cdot)$ is defined in (11).

The proof of the above proposition is provided in Appendix C. Now, we discuss the method of selecting the terminating time T^* . In this work, we consider the following two discrepancy functions:

- The Morozov's conventional discrepancy function:

$$\chi(T) = \frac{1}{\sqrt{2}} \|u_{im}(p^\delta(T))\|_{0,\Omega} - C_0 \tau \delta, \quad (50)$$

where $u = u_{re} + iu_{im}$ is the solution of (17) with noisy data, and τ is a fixed positive number.

- The total energy discrepancy function:

$$\chi_{TE}(T) = V(p^\delta(T)) + \frac{1}{2} \|\dot{p}^\delta(T)\|_P^2 - C_0^2 \tau^2 \delta^2, \quad (51)$$

where $V(p^\delta) = \frac{1}{2} \|u_{im}(p^\delta)\|_{0,\Omega}^2$.

Lemma 12. Under the assumption $\tau > 1$, the following two assertions hold.

- If $\|u_{im}(p_0)\|_{0,\Omega} \geq \sqrt{2} C_0 \tau \delta$, then $\chi(T)$ has at least one root.
- If $\|u_{im}(p_0)\|_{0,\Omega} + \|\dot{p}_0\|_P^2 \geq 2C_0^2 \tau^2 \delta^2$, then $\chi_{TE}(T)$ has a unique solution.

Proof. The continuity of $\chi(T)$ and $\chi_{TE}(T)$ is obviously according to Lemma 3 and Theorem 4. From Proposition 3 and the assumption of the lemma, we conclude that

$$\lim_{T \rightarrow \infty} \chi(T) \leq C_0(1 - \tau)\delta < 0 \text{ and } \lim_{T \rightarrow \infty} \chi_{TE}(T) \leq C_0^2(1 - \tau^2)\delta^2 < 0, \quad (52)$$

and $\chi(0) = \frac{1}{\sqrt{2}} \|u_{im}(p_0)\|_{0,\Omega} - C_0 \tau \delta > 0$ and $\chi_{TE}(0) = V(p_0) + \|\dot{p}_0\|_P^2 - C_0^2 \tau^2 \delta^2 > 0$, which implies the existence of the root of $\chi(T)$ and $\chi_{TE}(T)$.

The non-growing of $\chi_{TE}(T)$ is straightforward according to $\dot{\chi}_{TE} = -\eta \|\dot{p}^\delta\|_P^2$ for the fixed damping parameter and $\dot{\chi}_{TE} = -\frac{r}{T} \|\dot{p}^\delta\|_P^2$ for the dynamic damping parameter.

Finally, let us show that $\chi_{TE}(T)$ has a unique solution. We prove this by contradiction. Since $\chi_{TE}(T)$ is a non-increasing function, a number T_0 exists so that $\chi_{TE}(T) = 0$ for $T \in [T_0, T_0 + \varepsilon]$ with some positive $\varepsilon > 0$. This means that $\dot{\chi}_{TE}(T) = -\eta \|\dot{p}^\delta\|_P^2 \equiv 0$ (or $\dot{\chi}_{TE}(T) = -\frac{r}{T} \|\dot{p}^\delta\|_P^2 \equiv 0$) in $(T_0, T_0 + \varepsilon)$. Hence, $\ddot{p}^\delta \equiv 0$ in $(T_0, T_0 + \varepsilon)$. Using Eq. (9) we conclude that for all $T > T_0$: $p^\delta(T) \equiv p^\delta(T_0)$. Since $\chi_{TE}(T_0) = 0$, we obtain that $\chi_{TE}(T) \equiv 0$ for $T > T_0$, which implies that $\lim_{T \rightarrow \infty} \chi_{TE}(T) = 0$. This contradicts the fact in (52). \square

Remark 4. It should be noted that Lemma 12 may still hold in the case $\tau \leq 1$. In many situations, e.g. for our numerical examples in Section 6, a small value of τ offers a better result, provided the existence of the root of χ or χ_{TE} .

Theorem 13 (Convergence for Noisy Data). Let $p^\delta(t)$ be the dynamic solution of (15)–(17). Then, if the terminating time point T^* is selected as the root of $\chi(T)$ or $\chi_{TE}(T)$, $p^\delta(T^*(\delta))$ converges weakly to p^\dagger in P as $\delta \rightarrow 0$.

Proof. We use the technique from [31, Theorem 2.4]. Let $\{\delta_n\}$ be a sequence converging to 0 as $n \rightarrow \infty$, and let $\{g_1^{\delta_n}, g_2^{\delta_n}\}$ be a corresponding sequence of noisy data with $\|g_1^{\delta_n} - g_1\|_{0,r} \leq \delta_n$ and $\|g_2^{\delta_n} - g_2\|_{0,r} \leq \delta_n$. For a triple $(\delta_n, g_1^{\delta_n}, g_2^{\delta_n})$, denote by $T_n^* = T^*(\delta_n)$ the corresponding terminating time point determined from the generalized discrepancy principles $\chi(T) = 0$ or $\chi_{TE}(T) = 0$.

Two possible cases exist. (i) T_n^* has a finite accumulation point T^* . (ii) $T_n^* \rightarrow \infty$ as $\delta_n \rightarrow 0$. For the case (i), without loss of generality we can assume that $T_n^* = T^*$ for all $n \in \mathbf{N}$. Hence, from the definition of T_n^* it follows that

$$\|u_{im}(p^{\delta_n}(T_n^*))\|_{0,\Omega} \leq C_0 \tau \delta_n. \quad (53)$$

Since $p^{\delta_n}(T_n^*)$ depends continuously on $\{g_1^{\delta_n}, g_2^{\delta_n}\}$ when T_n^* is fixed, we have

$$p^{\delta_n}(T_n^*) \rightarrow p(\cdot, T^*), \quad \|u_{im}(p^{\delta_n}(T_n^*))\|_{0,\Omega} \rightarrow \|u_{im}(p(T^*))\|_{0,\Omega}. \quad n \rightarrow \infty, \quad (54)$$

where $p(\cdot, t)$ denotes the dynamic solution of (15)–(17) with noise-free data. Letting $n \rightarrow \infty$ in (53) yields $\|u_{im}(p(T^*))\|_{0,\Omega} = 0$. Thus, $p(T^*) = p^\dagger$, a solution of (1), and with (54) we obtain the strong convergence: $p(T_n^*) \rightarrow p^\dagger$ in P as $n \rightarrow \infty$.

Now, consider the case (ii). According to the continuity of $p^{\delta n}(t)$, for any positive ε_0 and T_n^* , there exists a point $T^* < T_n^*$ such that

$$\|p^{\delta n}(T_n^*) - p^{\delta n}(T^*)\|_P \leq \varepsilon_0. \quad (55)$$

On the other hand, for any $q(\cdot) \in P$,

$$\begin{aligned} |(p^{\delta n}(T_n^*) - p^\dagger, q)_P| &\leq \\ |(p^{\delta n}(T_n^*) - p^{\delta n}(T^*), q)_P| &+ |(p^{\delta n}(T^*) - p(T^*), q)_P| + |(p(T^*) - p^\dagger, q)_P|. \end{aligned}$$

By inequality (55) and the weak convergence of $p(t)$, one can fix T^* so large that both inequalities $|(p^{\delta n}(T_n^*) - p^{\delta n}(T^*), q)_P| \leq \varepsilon/3$ and $|(p(T^*) - p^\dagger, q)_P| \leq \varepsilon/3$ hold. Now that T^* is fixed, we can apply the result of case (i) to conclude that a positive number $n_1 = n_1(T^*)$ exists such that for any $n \geq n_1$: $|(p^{\delta n}(T^*) - p(T^*), q)_P| \leq \varepsilon/3$. Combine the above inequalities to obtain $|(p^{\delta n}(T_n^*) - p^\dagger, q)_P| \leq \varepsilon$ for all $n \geq n_1$. Since ε is arbitrary, we complete the proof. \square

5. Full discretization and a novel iterative regularization algorithm

5.1. Space discretization

Following [32], we discretize the bounded domain Ω by mesh \mathcal{T} using non-overlapping triangles/tetrahedrons $\{\Delta_\mu\}_{\mu=1}^M$. We associate the mesh \mathcal{T} with the mesh function $h(x)$, which is a piecewise-constant function such that $h(x) \equiv \ell(\Delta_\mu)$ for all $x \in \Delta_\mu$, where $\ell(\Delta_\mu)$ is the longest side of $\Delta_\mu \in \mathcal{T}$. Define the mesh scale size as $h := \max_{x \in \Omega} h(x)$. Let $r(\Delta_\mu)$ be the radius of the maximal circle/ball contained in the triangle/tetrahedron Δ_μ . We make the following shape regularity assumption for every element $\Delta_\mu \in \mathcal{T}$: $c_1 \leq \ell(\Delta_\mu) \leq c_2 r(\Delta_\mu)$, where c_1 and c_2 are two positive constants. Now, we introduce the finite element space

$$\Psi^h = \{v \in C(\Omega) : v \in \mathcal{P}_1(\Delta_\mu) \text{ for all } \Delta_\mu \in \mathcal{T}\}, \quad (56)$$

where $\mathcal{P}_1(\Delta_\mu)$ denotes the set of all linear continuous functions on Δ_μ .

Denote $\Psi^h := \Psi^h \oplus i\Psi^h$. Then, Ψ^h is a finite element subspace of $\mathbf{H}^1(\Omega)$, and the finite element approximation of the BVP (18) is as follows:

$$\text{find } u^h \in \Psi^h \text{ such that } a(u^h, \psi^h) = f^\delta(\psi^h), \quad \forall \psi^h \in \Psi^h. \quad (57)$$

The problem (57) admits a unique solution $u^h \in \Psi^h$ according to Lemma 2. Similar to those in [12], it is not difficult to derive the following a priori finite element error estimates.

Theorem 14. Let $u \in \mathbf{H}^1(\Omega)$ be the solution of the problem (18) and $u^h \in \Psi^h$ be the finite element solution of problem (57) respectively. Then, for any $p(t) \in L^2((t_0, \infty), P)$ and almost every $t > 0$

$$\|u^h(p(t)) - u(p(t))\|_{1,\Omega} \leq C(\Omega)h (\|p(t)\|_{0,\Omega_0} + \|g_1^\delta\|_{0,r} + \|g_2^\delta\|_{0,r}).$$

Note that, in this section, we set $t_0 = 0$ or 1, corresponding to the model (15) with different damping parameter $\eta(t) = \text{const.}$ or r/t . Now we are in a position to discretize the second order evolution equation (15). For this purpose, set $P^h = P \cap \Psi^h$ and the orthogonal projection operator $\Pi^h : P \rightarrow P^h$

$$(\Pi^h p, q^h)_{k,\Omega_0} = (p, q^h)_{k,\Omega_0}, \quad \forall p \in P, q^h \in P^h, k = 0, 1. \quad (58)$$

Then for all $p \in H^{k+1}(\Omega_0)$ [33, Theorem 10.3.8]:

$$\|\Pi^h p - p\|_{m,\Omega_0} \leq C(\Omega)h^{k+1-m}|p|_{k+1,\Omega_0}, \quad m = 0, 1. \quad (59)$$

Introduce a discrete optimization problem

$$\min_{p \in P^h} V_h(p) = \min_{p \in P^h} \frac{1}{2} \|u_{im}^h(p)\|_{0,\Omega}^2, \quad (60)$$

where $u^h = u_{re}^h + iu_{im}^h \in \Psi^h$ is the weak solution of the problem (57), and a semi-discretized second order flow

$$\begin{cases} \ddot{p}^{\delta,h}(x, t) + \eta(t)\dot{p}^{\delta,h}(x, t) + w_{im}^h(x, t) = 0, & x \in \Omega_0, t \in (t_0, \infty), \\ p^{\delta,h}(x, t_0) = p_0^h, \dot{p}^{\delta,h}(x, t_0) = \dot{p}_0^h, & x \in \Omega_0, \end{cases} \quad (61)$$

where p_0^h and \dot{p}_0^h are projections of p_0 and \dot{p}_0 in P^h , w^h is the finite element solution to the joint problem

$$\begin{cases} -\Delta w(x, t) + c w(x, t) = u_{im}^h(p^{\delta,h}(x, t)), & x \in \Omega_0, t \in (t_0, \infty), \\ \frac{\partial w(x, t)}{\partial \mathbf{n}} + i w(x, t) = 0, & x \in \Gamma, t \in (t_0, \infty), \end{cases} \quad (62)$$

and $u_{im}^h(p^{\delta,h}(x, t))$ is the imaginary part of the solution of (57), with p^δ replaced by $p^{\delta,h}$.

Proposition 4. Let $w^\delta \in \mathbf{H}^1(\Omega)$ be the weak solution of (18) with $p^\delta(t)$ replaced by $p^{\delta,h}(t)$, and $w^{\delta,h} \in \Psi^h$ be the finite element solution of (62). Then, a constant $C(\Omega)$ exists such that for any $p^{\delta,h}(t) \in L^2((t_0, \infty), P^h)$, and almost every $t \in [t_0, \infty)$,

$$\|w^{\delta,h}(p^{\delta,h}(t)) - w^\delta(p^{\delta,h}(t))\|_{1,\Omega} \leq C(\Omega)h (\|p^{\delta,h}(t)\|_{0,\Omega_0} + \|g_1^\delta\|_{0,r} + \|g_2^\delta\|_{0,r}).$$

Combining Theorems 4 and 14, Proposition 4, as well as the definition of Π^h , it is not difficult to obtain the following estimate.

Proposition 5. Let $p^\delta(t) \in P$ and $p^{\delta,h}(t) \in P^h$ be solutions of (15) and (61) respectively. Then, a constant $C(\Omega)$ exists such that for almost every $t \in [t_0, \infty)$,

$$\|p^{\delta,h}(t) - p^\delta(t)\|_P \leq C(\Omega)h (\|g_1^\delta\|_{0,r} + \|g_2^\delta\|_{0,r}).$$

Now, we present the main result in this subsection.

Theorem 15 (Convergence of the Finite Element Solution). Let $p^{\delta,h} \in P^h$ be solution of (61). Suppose that for almost every $t > 0$ and $\delta \geq 0$, $p^\delta(t) \in H^1(\Omega_0)$. Then, under the assumption of Theorem 13, we have the strong convergence, i.e., $p^{\delta,h}(T^*(\delta)) \rightarrow p^\dagger$ in P as $\delta, h \rightarrow 0$.

Proof. By the triangle inequality

$$\begin{aligned} & \|p^{\delta,h}(T^*(\delta)) - p^\dagger\|_P \\ & \leq \|p^{\delta,h}(T^*(\delta)) - p^\delta(T^*(\delta))\|_P + \|p^\delta(T^*(\delta)) - \Pi^h p^\delta(T^*(\delta))\|_P \\ & \quad + \|\Pi^h p^\delta(T^*(\delta)) - \Pi^h p^\dagger\|_P + \|\Pi^h p^\dagger - p^\dagger\|_P, \end{aligned}$$

it suffices to show the convergence of all terms in the right-hand side of the above inequality. The convergence of the first term follows from Proposition 5, while the second and fourth terms converge to 0 because of the inequality (59). Finally, the convergence of the third term follows from Theorem 13 and the assertion (b) of Remark 2. \square

Finally, we give a sketch of the finite element method for problems (16) and (17). For conciseness, by slightly abusing the notation, we rewrite $p^{\delta,h}$, $\dot{p}^{\delta,h}$ and $\ddot{p}^{\delta,h}$ to p^h , \dot{p}^h and \ddot{p}^h . Let m be the number of the nodes of triangulation \mathcal{T} , and $\{\psi_l\}_{l=1}^m$ be the nodal basis functions of the linear finite element space Ψ^h associated with the grid points $\{x_l\}_{l=1}^m$. Then $u^h(x, t) = \sum_{l=1}^m u_l(t)\psi_l(x)$ with $u_l(t) = u^h(x_l, t) \in L^2((t_0, \infty), \mathbb{C})$ and $w^h(x, t) = \sum_{l=1}^m w_l(t)\psi_l(x)$ with $w_l(t) = w^h(x_l, t) \in L^2((t_0, \infty), \mathbb{C})$. Denote $\{x_{k_l}\}_{l=1}^{m_0} = \{x_l\}_{l=1}^m \cap \overline{\Omega}_0$, $p^h(x, t) = \sum_{l=1}^{m_0} p_l(t)\psi_{k_l}(x)$ with $p_l(t) = p^h(x_{k_l}, t) \in L^2((t_0, \infty), \mathbb{R})$. As a result, the problem (57) reduces to the following algebraic system with any fixed t :

$$\begin{cases} (D + E)\mathbf{u}_{re}(t) - F\mathbf{u}_{im}(t) = B\mathbf{p}(t) + \mathbf{b}_2, \\ F\mathbf{u}_{re}(t) + (D + E)\mathbf{u}_{im}(t) = \mathbf{b}_1, \end{cases} \quad (63)$$

where

$$\begin{aligned} D &= [d_{ls}]_{m \times m}, d_{ls} = \int_{\Omega} \nabla \psi_s \cdot \nabla \psi_l dx, \quad E = [e_{ls}]_{m \times m}, e_{ls} = \int_{\Omega} \psi_s \psi_l dx, \\ F &= [f_{ls}]_{m \times m}, f_{ls} = \int_{\Gamma} \psi_s \psi_l ds, \quad B = [b_{lj}]_{m \times m_0}, b_{lj} = \int_{\Omega_0} \psi_l(x) \psi_{k_j}(y) dx, \\ \mathbf{b}_1 &= [b_{1,l}]_{m \times 1}, b_{1,l} = \int_{\Gamma} g_1^\delta \psi_l ds, \quad \mathbf{b}_2 = [b_{2,l}]_{m \times 1}, b_{2,l} = \int_{\Gamma} g_2^\delta \psi_l ds, \\ \mathbf{u}_{re} &= [u_{re,l}]_{m \times 1}, \mathbf{u}_{im} = [u_{im,l}]_{m \times 1}, \mathbf{p} = [p_j]_{m_0 \times 1}, l, s = \overline{1, m}, j = \overline{1, m_0}. \end{aligned}$$

Similarly, for any fixed t , finding a weak solution of (62) reduces to solve the following system of linear equations

$$\begin{cases} (D + E)\mathbf{w}_{re}(t) - F\mathbf{w}_{im}(t) = E\mathbf{u}_{im}(t), \\ F\mathbf{w}_{re}(t) + (D + E)\mathbf{w}_{im}(t) = \mathbf{0}. \end{cases} \quad (64)$$

5.2. Time discretization and a novel iterative regularization algorithm

The second order evolution equation (15) with an appropriate numerical discretization scheme for the artificial time variable yields a concrete second order iterative regularization method. The damped symplectic integrators are extremely attractive for solving second order systems, since the schemes are closely related to the canonical transformations [34], and the trajectories of the discretized second flow usually keep some intrinsic invariants of the system. In this paper, we use the Störmer–Verlet method, which belongs to the family of symplectic integrators.

Denote $q^h(x, t) = \dot{p}^h(x, t)$, and rewrite (61) into the first order system

$$\begin{cases} \dot{q}^h = -\eta q^h - w_{im}^h \chi_{\Omega_0}, \\ \dot{p}^h = q^h, \\ p^h(t_0) = p_0^h, q^h(t_0) = \dot{p}_0^h. \end{cases} \quad (65)$$

Apply the Störmer–Verlet method to the system (65) to obtain that at the k th iteration

$$\begin{cases} q_{k+\frac{1}{2}}^h = q_k^h - \frac{\Delta t}{2} \left(\eta(t_k) q_{k+\frac{1}{2}}^h + w_{im}^h(p_k^h) \chi_{\Omega_0} \right), \\ p_{k+1}^h = p_k^h + \Delta t q_{k+\frac{1}{2}}^h, \\ q_{k+1}^h = q_{k+\frac{1}{2}}^h - \frac{\Delta t}{2} \left(\eta(t_{k+1}) q_{k+\frac{1}{2}}^h + w_{im}^h(p_{k+1}^h) \chi_{\Omega_0} \right), \\ q^h(t_0) = \dot{p}_0^h, p^h(t_0) = p_0^h, \end{cases} \quad (66)$$

where $p_k^h = p^{\delta,h}(t_k)$, and Δt is the time step size.

Taking into account of the discrepancy principle for choosing the terminating time point, the newly developed numerical algorithm is proposed as follows:

Algorithm 1 The Störmer–Verlet based SOAR for inverse source problem (1).

Require: Boundary data $\{g_1^\delta, g_2^\delta\}$. Noise level δ . Damping parameter $\eta(t)$. Time step size Δt . The permissible region Ω_0 . Triangulation \mathcal{T} of domain Ω with the nodal basis functions $\{\psi_i\}_{i=1}^m$. Precision number ϵ_0 . Initial values: $(\mathbf{p}^0, \mathbf{q}^0)$. Iteration index: $k \leftarrow 0$.

Ensure: The estimated source term: $p^h = \sum_{l=1}^{m_0} \mathbf{p}_l^k \psi_{k_l}$.

- 1: **while** $\chi(t_k) > \epsilon_0$ or $\chi_{TE}(t_k) > \epsilon_0^2$ **do**
 - 2: Solve (63) and (64) with source \mathbf{p}^k to get \mathbf{w}_{im}^k .
 - 3: $\mathbf{q}^{k+\frac{1}{2}} \leftarrow \mathbf{q}^k - \frac{\Delta t}{2} \left(\eta(t_k) \mathbf{q}^{k+\frac{1}{2}} + \mathbf{w}_{im}^k \right)$
 - 4: $\mathbf{p}^{k+1} \leftarrow \mathbf{p}^k + \Delta t \mathbf{q}^{k+\frac{1}{2}}$
 - 5: Solve (63) and (64) with source \mathbf{p}^{k+1} to get \mathbf{w}_{im}^{k+1} .
 - 6: $\mathbf{q}^{k+1} \leftarrow \mathbf{q}^{k+\frac{1}{2}} - \frac{\Delta t}{2} \left(\eta(t_{k+1}) \mathbf{q}^{k+\frac{1}{2}} + \mathbf{w}_{im}^{k+1} \right)$
 - 7: $t_{k+1} \leftarrow t_k + \Delta t$
 - 8: $k \leftarrow k + 1$
 - 9: **end while**
-

6. Numerical results

We present numerical examples to demonstrate the effectiveness of the proposed second order asymptotical regularization (SOAR) methods. For the sake of simplicity, we refer to SOAR1 as Algorithm 1 when η is constant and χ is used; SOAR2 when η is constant and χ_{TE} is used; SOAR3 when $\eta = r/t$ and χ is used; SOAR4 when $\eta = r/t$ and χ_{TE} is used. All the computations were performed on a dual core personal computer of 8.00 GB RAM with MATLAB version R2007b.

6.1. Measurements and some parameters setting

Given the problem domain Ω and a true source function p^\dagger in $\Omega_0 \subset \Omega$, by using the standard linear finite element method described in Section 5.1, we solve the BVP

$$-\Delta u + u = p^\dagger \chi_{\Omega_0} \text{ in } \Omega, \text{ and } \frac{\partial u}{\partial \mathbf{n}} = g_2 = 0 \text{ on } \Gamma \quad (67)$$

on a quite fine mesh to get $u^h \in \Psi^h$. Then, the exact Dirichlet boundary data is given by $g_1 = u^h|_\Gamma$. In this section, we study the following three model problems, some of which are taken from [22]:

- **Example 1:** $\Omega = \{(x_1, x_2) \in \mathbb{R}^2 | x_1^2 + x_2^2 < 1\}$, $\Omega_0 = \{(x_1, x_2) \in \mathbb{R}^2 | -0.5 < x_1, x_2 < 0.5\}$. $p^\dagger(x_1, x_2) = (1 + x_1 + x_2) \chi_{\Omega_0}$. g_1 is computed on a mesh with mesh size $h = 0.01386$, 144 929 nodes and 288 768 elements;
- **Example 2:** Ω is the same as Example 1, $\Omega_0 = \Omega_1 \cup \Omega_2$ with $\Omega_1 = \{(x_1, x_2) \in \mathbb{R}^2 | (x_1 + 0.5)^2 + x_2^2 < 0.01\}$ and $\Omega_2 = \{(x_1, x_2) \in \mathbb{R}^2 | (x_1 - 0.5)^2 + x_2^2 < 0.01\}$. $p^\dagger(x_1, x_2) = (1 + x_1 + x_2) \chi_{\Omega_1} + e^{1+x_1+x_2} \chi_{\Omega_2}$. g_1 is computed on a mesh with $h = 0.01228$, 156 225 nodes and 311 296 elements;
- **Example 3:** $\Omega = \{(x_1, x_2, x_3) \in \mathbb{R}^3 | x_1^2 + x_2^2 + x_3^2 < 1\}$. $p^\dagger(x_1, x_2, x_3) = 10 + x_1 + x_2 + x_3$ in $\Omega_0 = \{(x_1, x_2, x_3) \in \Omega | (x_1 + 0.5)^2 + (x_2 + 0.5)^2 + (x_3 - 0.5)^2 \leq 0.04\}$. g_1 is computed on a mesh with $h = 0.1036$, 38 812 nodes and 220 636 elements.

The artificial noisy boundary data are generated according to

$$g_{1,2}^\delta(x) = [1 + \delta' \cdot (2 \text{rand}(x) - 1)] g_{1,2}(x)$$

for all $x \in \Gamma$, where $\text{rand}(x)$ denotes the random value from a uniform distribution on $[0, 1]$, and δ' represents the relative error level. The noise level of measurement data is calculated by $\delta = \max_{j=1,2} \|g_j^\delta - g_j\|_{\infty, \Gamma}$.

Table 1L2Err and IterN vs. τ with $\Delta t = 10$, $\eta = 0.01$ or $10/t$.

τ	SOAR1		SOAR2		SOAR3		SOAR4	
	L2Err	IterN	L2Err	IterN	L2Err	IterN	L2Err	IterN
Example 1								
0.00001	0.2540	N_{\max}	0.2540	N_{\max}	0.7086	N_{\max}	0.3337	1410
0.00005	0.2540	N_{\max}	0.2540	N_{\max}	0.5726	3233	0.3284	1386
0.0001	0.2540	N_{\max}	0.2540	N_{\max}	0.4244	1847	0.3123	1313
0.0005	0.1034	514	0.1022	504	0.1036	326	0.1084	338
0.001	0.0471	96	0.0495	118	0.0300	37	0.0299	33
0.005	0.0406	55	0.0438	77	0.0275	22	0.0275	22
0.01	0.0609	33	0.0502	68	0.0394	19	0.0394	19
0.05	0.1936	11	0.0609	33	0.0830	14	0.0830	14
0.1	0.1420	8	0.1656	24	0.0830	14	0.0830	14
Example 2								
0.00001	0.0532	540	0.0535	371	0.0527	320	0.0573	212
0.00005	0.0535	371	0.0582	201	0.0532	271	0.0575	210
0.0001	0.0552	275	0.0589	190	0.0551	236	0.0582	203
0.0005	0.0663	96	0.0632	129	0.0691	118	0.0680	124
0.001	0.0678	86	0.0647	113	0.0774	88	0.0778	87
0.005	0.0860	32	0.0679	82	0.1120	22	0.1150	20
0.01	0.0860	32	0.0730	68	0.1241	16	0.1278	15
0.05	0.1460	22	0.1343	36	0.1518	10	0.1653	11
0.1	0.2776	5	0.1460	22	0.1518	10	0.1518	10
Example 3								
0.00001	0.0646	N_{\max}	0.0646	N_{\max}	0.2696	N_{\max}	0.2380	4536
0.00005	0.0646	N_{\max}	0.0646	N_{\max}	0.2696	N_{\max}	0.0817	1413
0.0001	0.0646	N_{\max}	0.0562	164	0.2696	N_{\max}	0.0555	187
0.0005	0.0726	17	0.0604	92	0.0740	29	0.0729	30
0.001	0.0726	17	0.0719	74	0.0756	28	0.0756	28
0.005	0.0706	6	0.1023	44	0.0945	24	0.1314	21
0.01	0.0706	6	0.1678	30	0.1163	22	0.2193	17
0.05	0.2268	5	0.9577	1	0.2927	3	0.9577	1
0.1	0.6655	3	0.9577	1	0.5304	2	0.9577	1

For the implementation of Algorithm 1, in all experiments below, we set $t_0 = 1$ and $\epsilon_0 = 10^{-4}$. Due to the assumptions of Lemma 12 and Theorem 13, theoretically, $\tau > 1$ (used in (50) and (51)), and $\|u_{im}(p_0)\|_{0,\Omega} \geq \sqrt{2}C_0\tau\delta$ or $V(p_0) + \frac{1}{2}\|\dot{p}_0\|_P^2 \geq C_0^2\tau\delta^2$ are required for the convergence. However, as mentioned in Remark 4 and as indicated by our numerical experiments, these two requirements are just sufficient conditions. In fact, small value of τ , i.e. $\tau < 1$, produces better behaviors in both solution accuracy and convergence rate, see Table 1. For the choice of the initial data (p_0, \dot{p}_0) for SOAR, according to the numerical experiments (for the concision of the statement, we omit the related numerical results), in most cases, the initial data (p_0, \dot{p}_0) does not affect the quality in solutions (the value of “L2Err”), but may influence the algorithm speed (the value of “IterN”). The closer the initial data (p_0, \dot{p}_0) is to the unknown exact solution, the less of the iteration number required is. Without knowledge of the exact solution, in this section, we simply set $p_0 = \dot{p}_0 = 0$. Moreover, to assess the accuracy of the approximate solutions, we define the L^2 -norm relative error for an approximate solution p^h : $L2Err := \|p^h - p^\dagger\|_P / \|p^\dagger\|_P$.

In the following, for Example 1, the approximate sources are reconstructed over a mesh with mesh size $h = 0.0774$, 2325 nodes and 4512 elements. For Example 2, the approximate sources are reconstructed over a mesh with mesh size $h = 0.0678$, 2505 nodes and 4864 elements. For Example 3, in Section 6.2, the approximate sources are reconstructed over a mesh with mesh size $h = 0.2719$, 2413 nodes, 12 372 elements, while in Sections 6.3 and 6.4, the approximate sources are reconstructed over a mesh with mesh size $h = 0.2238$, 4121 nodes, 21 841 elements. We use $N_{\max} := 5000$ as the maximal number of iterations where the iteration (66) stops in all of simulations. It mainly means that the used parameters for our methods (66) are inefficient. If the algorithm (66) with some inappropriate parameters diverges, i.e. the L^2 -norm relative error “L2Err” is significantly increasing after a certain step of iterations, we will use the mark “div”.

6.2. 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 Δt , damping parameter η when it is constant or r when $\eta(t) = r/t$, and thus to give a guide on the choices of them in practice. For focusing on the effect of these parameters on Algorithm 1, we fix $\delta' = 5\%$ in this subsection.

We first investigate the influence of parameter τ on the solution accuracy and convergence rate. For this purpose, we additionally set $\Delta t = 10$, $\eta = 0.01$ when η is constant or $\eta = 10/t$ when η is dynamic. The detailed L^2 -norm relative

Table 2L2Err and IterN vs. Δt with $\eta = 0.01$ or $10/t$, $\tau = 0.005$ for Example 1; $\tau = 0.001$ for Example 2; $\tau = 0.0005$ for Example 3.

Δt	SOAR1		SOAR2		SOAR3		SOAR4	
	L2Err	IterN	L2Err	IterN	L2Err	IterN	L2Err	IterN
Example 1								
0.5	0.0282	127	0.0480	1833	0.0615	442	0.0615	442
1	0.0278	172	0.0480	915	0.0615	221	0.0615	221
5	0.0414	120	0.0478	179	0.0608	42	0.0608	42
10	0.0406	55	0.0438	77	0.0275	22	0.0275	22
15	div	–	div	–	div	–	div	–
Example 2								
0.5	0.0720	1167	0.0646	2318	0.0775	1800	0.0778	1784
1	0.0720	583	0.0646	1158	0.0775	900	0.0778	892
5	0.0693	143	0.0646	231	0.0773	180	0.0777	178
10	0.0678	86	0.0647	113	0.0774	88	0.0778	87
15	0.0673	60	0.0651	73	0.0771	57	0.0776	56
20	0.0659	50	0.0659	50	0.0757	42	0.0757	42
25	div	–	div	–	div	–	div	–
Example 3								
0.5	0.0739	123	0.0598	1860	0.0733	560	0.0771	721
1	0.0708	283	0.0599	930	0.0732	280	0.0771	360
5	0.0672	122	0.0600	185	0.0739	57	0.0765	61
20	0.0653	35	0.0609	45	0.0725	25	0.0745	24
40	0.0655	16	0.0635	21	0.0673	10	0.0660	15
60	0.0648	10	0.0628	11	0.0539	22	0.0495	21
65	div	–	div	–	div	–	div	–

errors ‘L2Err’ and the corresponding iterative numbers ‘IterN’ for different values of τ are shown in Table 1, which shows that, on one hand, the smaller τ is, on the whole, the better the solution accuracy is; on the other hand, the smaller τ is, the more the iterative number for stopping Algorithm 1 is. It is no surprise that the parameter τ does not involve the computation of the approximate solutions itself. It is used in stop criterion and only affects the iterative number where Algorithm 1 stops. Therefore, it is natural that a larger iterative number produces a better approximate solution, and this also confirms the asymptotical behavior of the proposed method. Generally, $\tau < 1$ is enough to produce reasonable approximate solutions. Moreover, we can also see from Table 1 that for small τ , SOAR2 and SOAR4 behave much better with respect to both the solution accuracy and iterative number than SOAR1 and SOAR1 which indicates the merits of using χ_{TE} . For the trade-off of the solution accuracy and the iterative number, as indicated in Table 1, in the following experiments, we set $\tau = 0.005$ for Example 1; $\tau = 0.001$ for Example 2; $\tau = 0.0005$ for Example 3.

Next we investigate the influence of time step size Δt on the values of “L2Err” and “IterN”. To this end, set $\eta = 0.01$ or $10/t$, $\tau = 0.005$ for Example 1; $\tau = 0.001$ for Example 2; $\tau = 0.0005$ for Example 3. The results are given in Table 2, which shows that the bigger the time step size Δt is, the faster the iteration is. However, Δt should not be too big. Otherwise, the iteration will blow up as it breaks the consistency of the numerical scheme. According to Table 2, in the remaining experiments, we choose $\Delta t = 10$ for Example 1; $\Delta t = 20$ for Example 2; $\Delta t = 60$ for Example 3.

Finally, we discuss the influence of the damping parameter η on the solution accuracy and the convergence rate. In the experiments, set $\tau = 0.005$, $\Delta t = 10$ for Example 1; $\tau = 0.001$, $\Delta t = 20$ for Example 2; $\tau = 0.0005$, $\Delta t = 60$ for Example 3. The results are displayed in Table 3. We conclude from Table 3 that for constant η , $\eta \leq 0.1$ can lead to reasonable approximate solutions for four cases of Algorithm 1. Nevertheless, η should not be too small. Too small η brings oscillation in solution accuracy. Table 3 also shows that like constant η , for dynamic damping parameter $\eta = r/t$, the factor r should be neither too small nor too big. Too small r also brings oscillation in solution accuracy. Therefore, in the remaining experiments, set $\eta = 0.05$ or $10/r$ for Example 1; $\eta = 0.005$ or $15/r$ for Example 2; $\eta = 0.005$ or $10/r$ for Example 3.

6.3. The convergence of approximate solutions with respect to the noise level

We validate in this subsection the convergence result of Theorem 13. To the end, as indicated by Section 6.2, we set $\tau = 0.005$, $\Delta t = 10$, $\eta = 0.05$ or $10/t$ in Example 1; $\tau = 0.001$, $\Delta t = 20$, $\eta = 0.005$ or $15/t$ in Example 2; $\tau = 0.0005$, $\Delta t = 60$, $\eta = 0.005$ or $10/t$ in Example 3, and implement Algorithm 1 for $\delta' = 0.2, 0.1, 0.05, 0.01, 0.005, 0.001$. The detailed errors and the corresponding iterative numbers are given in Table 4, from which we can see that in all examples and for all methods: SOAR1-SOAR4, on the whole, the solution accuracy improves when the noisy level decreases and, thus all methods are convergent.

Table 3

L2Err and IterN vs. $\eta = \text{const.}$ or r/t with $\tau = 0.005, \Delta t = 10$ in Example 1; $\tau = 0.001, \Delta t = 20$ in Example 2; $\tau = 0.0005, \Delta t = 60$ in Example 3.

η	Example 1				Example 2				Example 3			
	SOAR1		SOAR2		SOAR1		SOAR2		SOAR1		SOAR2	
	L2Err	IterN	L2Err	IterN	L2Err	IterN	L2Err	IterN	L2Err	IterN	L2Err	IterN
0.001	0.2080	207	0.3066	830	0.0856	343	0.0935	507	0.0555	44	0.0737	129
0.005	0.0676	63	0.0758	164	0.0531	89	0.0529	103	0.0497	22	0.0483	26
0.01	0.0406	55	0.0438	77	0.0659	50	0.0659	50	0.0648	14	0.0628	11
0.05	0.0423	14	0.0423	14	0.0771	100	0.0794	88	0.0736	9	0.0736	9
0.1	0.0614	26	0.0687	24	0.0771	199	0.0794	175	div	–	div	–
0.5	div	–	div	–	div	–	div	–	div	–	div	–
r	Example 1				Example 2				Example 3			
	SOAR3		SOAR4		SOAR3		SOAR4		SOAR3		SOAR4	
	L2Err	IterN	L2Err	IterN	L2Err	IterN	L2Err	IterN	L2Err	IterN	L2Err	IterN
1	0.0455	40	0.3481	610	0.0982	264	1.0044	N_{max}	0.0815	87	0.3964	443
5	0.0484	16	0.0376	17	0.0767	33	0.0730	37	0.0695	8	0.0695	8
10	0.0275	22	0.0275	22	0.0757	42	0.0757	42	0.0539	22	0.0495	21
15	0.0651	39	0.0619	41	0.0705	83	0.0669	85	0.0836	57	0.1202	50
20	0.4062	143	0.4303	137	0.1595	261	0.1543	264	0.6134	105	0.6254	97

Table 4

L2Err and IterN vs. δ' with $\tau = 0.005, \Delta t = 10, \eta = 0.05$ or $10/t$ in Example 1; $\tau = 0.001, \Delta t = 20, \eta = 0.005$ or $15/t$ in Example 2; $\tau = 0.0005, \Delta t = 60, \eta = 0.005$ or $10/t$ in Example 3.

δ'	SOAR1		SOAR2		SOAR3		SOAR4	
	L2Err	IterN	L2Err	IterN	L2Err	IterN	L2Err	IterN
Example 1								
0.2	0.0420	10	0.0355	13	0.1056	18	0.1056	18
0.1	0.0732	10	0.0371	14	0.0491	19	0.0491	19
0.05	0.0423	14	0.0335	17	0.0275	22	0.0275	22
0.01	0.0247	18	0.0245	22	0.0265	26	0.0248	29
0.005	0.0252	18	0.0245	24	0.0245	30	0.0245	30
0.001	0.0245	22	0.0244	25	0.0244	30	0.0245	33
Example 2								
0.2	0.1391	68	0.1434	75	0.1388	38	0.1488	38
0.1	0.0933	82	0.0946	89	0.1384	78	0.1384	78
0.05	0.0531	89	0.0529	103	0.0705	83	0.0669	85
0.01	0.0439	103	0.0433	117	0.0495	94	0.0476	96
0.005	0.0458	110	0.0447	121	0.0469	94	0.0462	96
0.001	0.0450	117	0.0445	121	0.0465	96	0.0462	97
Example 3								
0.2	0.0645	19	0.0621	22	0.0641	19	0.0641	19
0.1	0.0588	19	0.0600	26	0.0583	19	0.0623	22
0.05	0.0636	19	0.0640	30	0.0664	23	0.0664	23
0.01	0.0594	26	0.0582	33	0.0605	26	0.0605	26
0.005	0.0606	26	0.0595	33	0.0618	26	0.0609	29
0.001	0.0604	26	0.0592	33	0.0609	30	0.0606	29

6.4. Comparison with other methods

In this subsection, we compare the behaviors regarding the solution accuracy and the convergence rate between SOAR and three existing methods; that is, the Nesterov's method, the ν -method and the dynamical regularization method (DRM) proposed in [22]. Recall that we use \mathbf{p} as the coefficients of the finite element solution p^h , see Algorithm 1 for the detail. In all methods, we set $\tau = 0.005$ for Example 1; $\tau = 0.001$ for Example 2; $\tau = 0.0005$ for Example 3, $\mathbf{p}^0 = \mathbf{0}, \mathbf{q}^0 = \mathbf{0}$ if \mathbf{q} is involved, and $\mathbf{p}^1 = \mathbf{p}^0$ if the method is a two-step one. Moreover, in SOAR2 and SOAR4, the total energy discrepancy principle χ_{TE} is used, while, in all other methods, the usual discrepancy function χ is used.

For methods SOAR1-SOAR4, set $\Delta t = 10, \eta = 0.05$ or $10/t$ in Example 1; $\Delta t = 20, \eta = 0.005$ or $15/t$ in Example 2; $\Delta t = 60, \eta = 0.005$ or $10/t$ in Example 3. We remark that on the one hand, these chosen parameters are not the optimal ones; on the other hand, a large range of values of these parameters could produce satisfactory approximate sources p^h .

For the inverse source problem (1) with CCMB formulation, DRM yields the following iteration

$$\begin{cases} \mathbf{q}^{k+1} = \frac{1}{1+\eta\Delta t} \mathbf{q}^k - \frac{\Delta t}{1+\eta\Delta t} (\mathbf{w}_{im}^k + \varepsilon(t_k) \mathbf{p}^k), \\ \mathbf{p}^{k+1} = \mathbf{p}^k + \Delta t \mathbf{q}^{k+1}, \end{cases} \quad k = 0, 1, \dots, \quad (68)$$

Table 5

Comparison with the state-of-the-art methods. The CPU time is measured in seconds.

δ'	1%		5%		10%	
Example 1						
Methods	L2Err	IterN(CPU)	L2Err	IterN(CPU)	L2Err	IterN(CPU)
DRM	0.0259	463(52.22)	0.0613	278(32.03)	0.0924	192(21.84)
ν	0.0085	54(6.30)	0.0522	50(6.27)	0.0854	45(5.56)
Nesterov	0.0249	75(8.69)	0.0509	36(4.33)	0.0880	31(3.72)
SOAR1	0.0247	18(3.05)	0.0423	14(2.28)	0.0732	10(1.63)
SOAR2	0.0245	22(2.75)	0.0478	17(2.23)	0.0371	14(1.91)
SOAR3	0.0265	26(3.56)	0.0275	22(2.94)	0.0491	19(2.69)
SOAR4	0.0248	29(7.92)	0.0275	22(5.88)	0.0491	19(5.19)
Example 2						
DRM	0.0613	$N_{\max}(600.11)$	0.0771	2017(244.75)	0.0849	1247(149.88)
ν	0.0505	564(66.48)	0.0679	440(51.70)	0.0785	247(29.23)
Nesterov	0.0507	397(47.98)	0.0768	165(20.95)	0.0791	162(19.67)
SOAR1	0.0439	103(29.52)	0.0531	89(25.16)	0.0933	82(25.58)
SOAR2	0.0433	117(17.48)	0.0529	103(15.56)	0.0946	89(15.86)
SOAR3	0.0495	93(12.80)	0.0705	83(11.86)	0.1384	78(11.27)
SOAR4	0.0476	96(13.45)	0.0669	85(12.27)	0.1384	78(13.22)
Example 3						
DRM	0.0584	1947(1696.66)	0.0723	134(125.25)	0.0717	108(101.19)
ν	0.0509	491(446.11)	0.0649	218(196.13)	0.0695	54(56.34)
Nesterov	0.0533	349(317.91)	0.0675	177(163.80)	0.0745	39(45.63)
SOAR1	0.0594	26(136.02)	0.0636	19(111.27)	0.0588	19(112.14)
SOAR2	0.0582	33(143.56)	0.0640	30(128.84)	0.0600	30(128.84)
SOAR3	0.0605	26(117.91)	0.0664	23(106.38)	0.0583	19(92.69)
SOAR4	0.0605	26(115.64)	0.0664	23(107.28)	0.0623	22(101.91)

where $(\mathbf{w}_{re}^k, \mathbf{w}_{im}^k)$ solves (64) with \mathbf{u}_{im} replaced by \mathbf{u}_{im}^k , and $(\mathbf{u}_{re}^k, \mathbf{u}_{im}^k)$ solves (63) with \mathbf{p} replaced by \mathbf{p}^k . As suggested by numerical experiments of [22], we set $\eta = 1$, $\Delta t = 10$ for Example 1; $\Delta t = 20$ for Example 2; $\Delta t = 60$ for Example 3 and the regularization parameter $\varepsilon(t) = 0.1/(t \ln(t))$. It should be mentioned that DRM is not an acceleration method.

For the ν -method, it is defined as [15, § 6.3]

$$\mathbf{p}^{k+1} = \mathbf{p}^k + \mu_k(\mathbf{p}^k - \mathbf{p}^{k-1}) - \omega_k \mathbf{w}_{im}^k, \quad k = 1, 2, \dots \quad (69)$$

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

$$\mu_k = \frac{(k-1)(2k-3)(2k+2\nu-1)}{(k+2\nu-1)(2k+4\nu-1)(2k+2\nu-3)}, \quad \omega_k = 4 \frac{(2k+2\nu-1)(k+\nu-1)}{(k+2\nu-1)(2k+4\nu-1)}.$$

Note that \mathbf{w}_{im}^k in (69) has the same meaning as that in (68). We select the Chebyshev method as our special ν -method, i.e., $\nu = 1/2$.

The Nesterov's method is defined by [17]

$$\begin{cases} \mathbf{z}_k = \mathbf{p}^k + \frac{k-1}{k+\alpha-1}(\mathbf{p}^k - \mathbf{p}^{k-1}), & k = 1, 2, \dots, \\ \mathbf{p}^{k+1} = \mathbf{z}_k - \omega \mathbf{w}_{im}^k, \end{cases} \quad (70)$$

where $\alpha \geq 3$, \mathbf{w}_{im}^k has the same definition as that in (68) and (69). We apply (70) to Examples 1–3 with parameters $\alpha = 3$ and $\omega = 10$.

The results of the simulations are presented in Table 5, from which we conclude that, with properly chosen parameters, all the mentioned methods are stable and can produce satisfactory solutions. Compared with the dynamical regularization method, all other methods offer solutions with similar accuracy, but require considerably fewer iterations. Particularly, SOAR1–SOAR4 can be compared to the well-known Nesterov's method and the ν -method, in both the solution accuracy and the convergence rate. On the whole, for all examples, the total energy discrepancy function χ_{TE} leads to more accurate solution than the conventional discrepancy function χ , but with slightly more iterative numbers.

We finally plot the exact and some recovered sources computed with different methods for $\delta' = 5\%$ in Figs. 1–3 for Examples 1–3 respectively. For the conciseness of the paper, we omit the figures corresponding to $\delta' = 1\%$ and 10% .

7. Conclusions

This paper is devoted to developing Second Order Asymptotical Regularization (SOAR) methods for solving inverse source problems of elliptic partial differential equations given Dirichlet and Neumann boundary data. We show the convergence results of SOAR for both fixed and dynamic damping parameters. A symplectic scheme is applied for the numerical implementation of SOAR. This scheme yields a novel iterative regularization method. As shown by the numerical

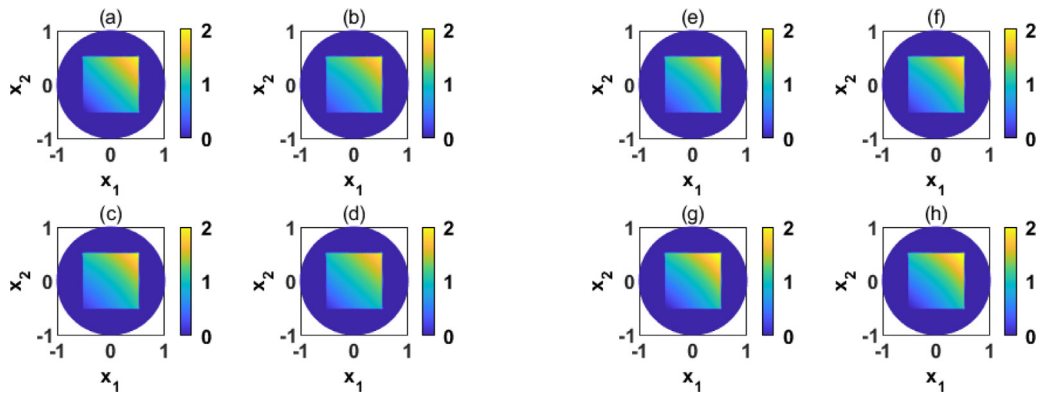


Fig. 1. The true and approximate sources. (a): p^+ ; (b): p^h by DRM; (c): p^h by Nesterov's method (d): p^h by ν -method; (e): p^h by SOAR1; (f): p^h by SOAR2; (g): p^h by SOAR3; (h): p^h by SOAR4.

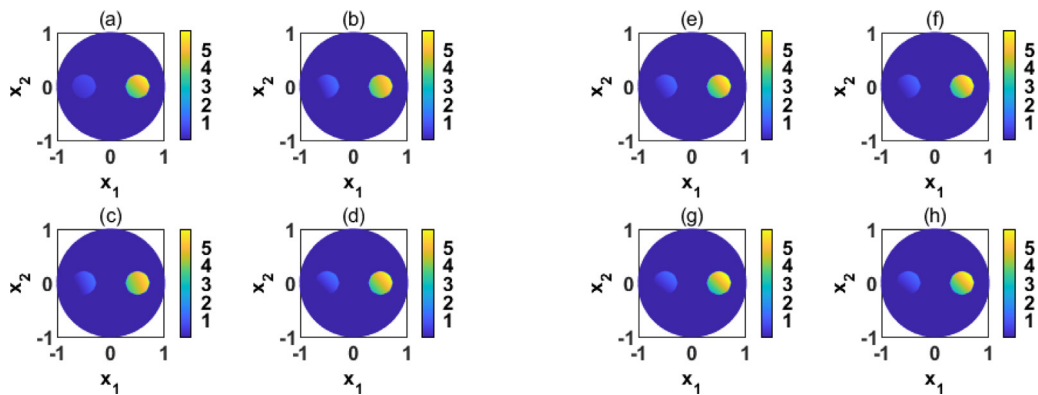


Fig. 2. The true and approximate sources. (a): p^+ ; (b): p^h by DRM; (c): p^h by Nesterov's method (d): p^h by ν -method; (e): p^h by SOAR1; (f): p^h by SOAR2; (g): p^h by SOAR3; (h): p^h by SOAR4.

results, the proposed SOAR methods are comparable to the Nesterov's acceleration method and the ν -method about the convergence rate. Moreover, in this paper, a conventional Morozov's discrepancy principle and a new total energy discrepancy principle are used for the stop criterion. Numerical experiments demonstrate that, in most cases, the newly developed total energy discrepancy principle works slightly better than the conventional Morozov's discrepancy principle. Similar to the Nesterov's acceleration method, the introduced SOAR can also be used to solve to non-linear ill-posed problems in partial differential equations, which will be one of the topics of our future work.

Acknowledgments

The authors thank the anonymous referees for their many constructive comments which have led to a significant improvement of the quality and the presentation of the paper. The work of Y. Zhang is supported by the Alexander von Humboldt foundation through a postdoctoral researcher fellowship. The work of R. Gong is supported by the National Natural Science Foundation of China (Nos. 11401304 and 11971230) and the Fundamental Research Funds for the Central Universities (No. NS2018047).

Appendix A. Proof of Theorem 4

Denote $q^\delta = \dot{p}^\delta$, $q^\delta(0) = \dot{p}^\delta(x, 0)$, and rewrite (15) as

$$\begin{cases} \dot{p}^\delta = q^\delta, \\ \dot{q}^\delta = -\eta q^\delta - w_{im} \chi_{\Omega_0}, \\ p^\delta(0) = p_0, q^\delta(0) = \dot{p}_0. \end{cases} \quad (\text{A.1})$$

By inequality (21) in Lemma 3, $w_{im} \chi_{\Omega_0}$ is continuously dependent on the source term p , hence, by the Cauchy–Lipschitz theorem, the first order nonautonomous system (A.1) has a unique global solution for the given initial data (p_0, \dot{p}_0) .

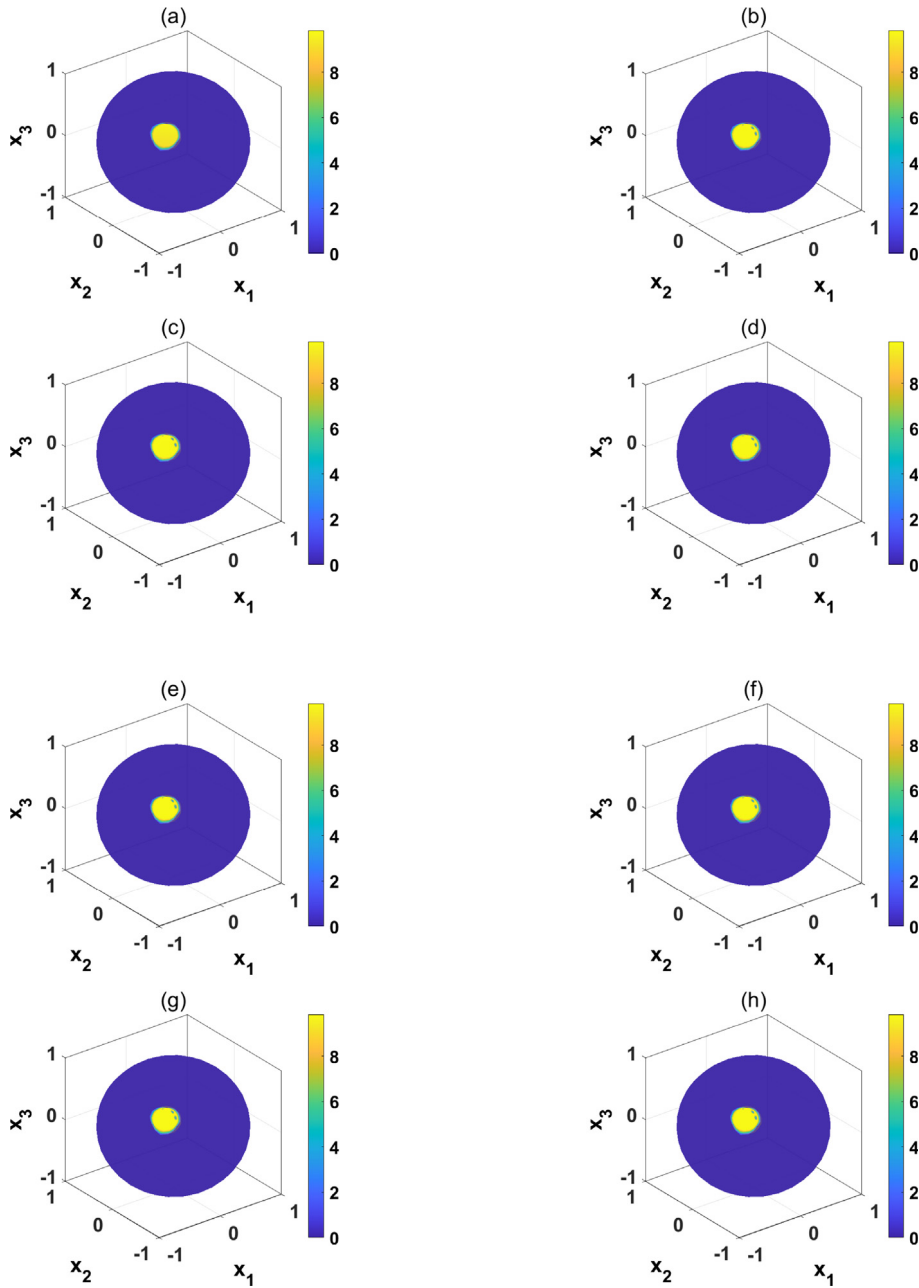


Fig. 3. The true and approximate sources. (a): p^δ ; (b): p^h by DRM; (c): p^h by Nesterov's method; (d): p^h by v -method; (e): p^h by SOAR1; (f): p^h by SOAR2; (g): p^h by SOAR3; (h): p^h by SOAR4.

Furthermore, by the standard arguments in elliptic PDEs theory [12,32], the global existence of the source function $p^\delta(x, t)$ implies the existence and uniqueness of the elliptic PDEs (16) and (17), which completes the proof of the global existence and uniqueness of the systems (15)–(17).

Now, we show the continuity of the solution p^δ with respect to the boundary data.

For any fixed t , define an operator $\mathcal{A} : P \rightarrow \mathbf{H}^1(\Omega)$ through $\mathcal{A}p(\cdot, t) = \hat{u}(\cdot, t)$ with $\hat{u}(\cdot, t) \in \mathbf{H}^1(\Omega)$ being the weak solution of

$$\begin{cases} -\Delta \hat{u}(x, t) + \hat{u}(x, t) = p(x, t)\chi_{\Omega_0}, & x \in \Omega, \quad t \in (0, \infty), \\ \frac{\partial \hat{u}(x, t)}{\partial \mathbf{n}} + i\hat{u}(x, t) = 0, & x \in \Gamma, \quad t \in (0, \infty). \end{cases}$$

Denote by $g = g_2 + ig_1$. For any $g \in \mathbf{L}^2(\Gamma)$, define operator $\mathcal{B} : \mathbf{L}^2(\Gamma) \rightarrow \mathbf{H}^1(\Omega)$ through $\mathcal{B}g = \tilde{u}$, where $\tilde{u} \in \mathbf{H}^1(\Omega)$ solves

$$\begin{cases} -\Delta \tilde{u}(x) + \tilde{u}(x) = 0, & x \in \Omega, \\ \frac{\partial \tilde{u}(x)}{\partial \mathbf{n}} + i\tilde{u}(x) = g, & x \in \Gamma. \end{cases}$$

Furthermore, for any $v \in \mathbf{H}^1(\Omega)$, we define $I_m : \mathbf{H}^1(\Omega) \rightarrow H^1(\Omega)$ through $I_mv = v_{im}$. Following standard arguments in the classical PDEs theory, all of \mathcal{A} , \mathcal{B} and I_m are bounded in the corresponding spaces. One the other hand, if we denote $g^\delta = g_2^\delta + ig_1^\delta$, we have

$$w_{im} = I_mv = I_m \mathcal{A} I_m (\mathcal{A} p^\delta + \mathcal{B} g^\delta) =: \mathcal{M} p^\delta + \mathcal{N} g^\delta.$$

Substitute the above equation into (15) to obtain

$$\begin{cases} \ddot{p}^\delta(x, t) + \eta \dot{p}^\delta(x, t) + \mathcal{M} p^\delta(x, t) = -\mathcal{N} g^\delta, & x \in \Omega_0, t \in (0, \infty), \\ p^\delta(x, 0) = p_0, \dot{p}^\delta(x, 0) = \dot{p}_0, & x \in \Omega_0. \end{cases}$$

If we define $\delta p = p^\delta - p$, it solves

$$\begin{cases} \ddot{\delta p}(x, t) + \eta \dot{\delta p}(x, t) + \mathcal{M} \delta p(x, t) = -\mathcal{N}(g^\delta - g), & x \in \Omega_0, t \in (0, \infty), \\ \delta p(x, 0) = \dot{\delta p}(x, 0) = 0, & x \in \Omega_0, \end{cases}$$

Applying the Cauchy–Lipschitz theorem again to deduce that for any fixed t , $\delta p(\cdot, t) \rightarrow 0$ in P when $g^\delta \rightarrow g$ in $\mathbf{L}^2(\Gamma)$. Consequently, $p^\delta(\cdot, t) \rightarrow p(\cdot, t)$ in P as $\delta \rightarrow 0$.

Appendix B. Proof of Lemma 5

Consider for every $t \in [0, \infty)$ the function $e(t) = e(t; p^\dagger) = \frac{1}{2} \|p(t) - p^\dagger\|_P^2$, where p^\dagger is defined in (13). Since $\dot{e}(t) = (p(t) - p^\dagger, \dot{p}(t))_P$ and $\ddot{e}(t) = \|\dot{p}(t)\|_P^2 + (p(t) - p^\dagger, \ddot{p}(t))_P$ for every $t \in [0, \infty)$, taking into account (9), we get

$$\ddot{e}(t) + \eta \dot{e}(t) + (p(t) - p^\dagger, w_{im}(p(t)))_P = \|\dot{p}(t)\|_P^2. \quad (\text{B.1})$$

Here, and later on, we denote $(p, u)_P = \int_{\Omega_0} p u dx$ for $P = L^2(\Omega_0)$ and $(p, u)_P = \int_{\Omega_0} p u dx + \int_{\Omega_0} \nabla p \cdot \nabla u dx$ for $P = H^1(\Omega_0)$. Moreover, $\|u\|_P = \sqrt{(u, u)_P}$.

On the other hand, by the convexity inequality of the functional $\|w_{im}(\cdot)\|_P^2$, we derive together with $w_{im}(p^\dagger) \equiv 0$ that

$$\|w_{im}(p(t))\|_P^2 = \|w_{im}(p(t))\|_P^2 - \|w_{im}(p^\dagger)\|_P^2 \leq (p(t) - p^\dagger, w_{im}(p(t)))_P. \quad (\text{B.2})$$

Combine (B.1) and the above inequality to obtain

$$\ddot{e}(t) + \eta \dot{e}(t) + \|w_{im}(p(t))\|_P^2 \leq \|\dot{p}(t)\|_P^2 \quad (\text{B.3})$$

or, equivalently (by using Eq. (15)),

$$\ddot{e}(t) + \eta \dot{e}(t) + \eta \frac{d\|\dot{p}(t)\|_P^2}{dt} + (\eta^2 - 1) \|\dot{p}(t)\|_P^2 + \|\ddot{p}\|_P^2 \leq 0. \quad (\text{B.4})$$

By the assumption $\eta \geq 1$, we deduce that

$$\ddot{e}(t) + \eta \dot{e}(t) + \eta \frac{d\|\dot{p}(t)\|_P^2}{dt} \leq 0, \quad (\text{B.5})$$

which means that the function $t \mapsto \dot{e}(t) + \eta e(t) + \eta \|\dot{p}(t)\|_P^2$ is monotonically decreasing. Hence a real number C exists such that

$$\dot{e}(t) + \eta e(t) + \eta \|\dot{p}(t)\|_P^2 \leq C, \quad (\text{B.6})$$

which implies $\dot{e}(s) + \eta e(s) \leq C$. By multiplying this inequality with $e^{\eta s}$ and then integrating from 0 to t , we obtain the inequality

$$e(t) \leq e(0)e^{-\eta t} + C(1 - e^{-\eta t})/\eta \leq e(0) + C/\eta.$$

Hence, $e(\cdot)$ is uniform bounded, and, consequently, $p(\cdot) \in L^\infty([0, \infty), P)$.

Now, consider the long-term behavior of \dot{p} . Define the Lyapunov function of the differential equation (15) by $\mathcal{E}(t) = V(p(t)) + \frac{1}{2} \|\dot{p}(t)\|_P^2$. It is not difficult to show that

$$\dot{\mathcal{E}}(t) = -\eta \|\dot{p}(t)\|_P^2 \quad (\text{B.7})$$

by looking at Eq. (15) and the differentiation of the energy function $\dot{\mathcal{E}}(t) = (\dot{p}(t), \ddot{p}(t) - u_{im}(p(t)))_P$. Hence, $\mathcal{E}(t)$ is non-increasing, and consequently, $\|\dot{p}(t)\|_P^2 \leq 2\mathcal{E}(0)$. Therefore, $\dot{p}(\cdot) \in L^\infty([0, \infty), P)$. Integrating both sides in (B.7),

we obtain

$$\int_0^\infty \|\dot{p}(t)\|_P^2 dt \leq \mathcal{E}(0)/\eta < \infty,$$

which yields $\dot{p}(\cdot) \in L^2([0, \infty), P)$ (and $\lim_{t \rightarrow \infty} \dot{p}(t) = 0$ since $\dot{p}(\cdot) \in L^\infty([0, \infty), P) \cap L^2([0, \infty), P)$).

Define

$$h(t) = \frac{\eta}{2} \|p(t) - p^\dagger\|_P^2 + (\dot{p}(t), p(t) - p^\dagger)_P. \quad (\text{B.8})$$

By elementary calculations, we derive that

$$\begin{aligned} \dot{h}(t) &= \eta(\dot{p}(t), p(t) - p^\dagger)_P + (\ddot{p}(t), p(t) - p^\dagger)_P + \|\dot{p}(t)\|_P^2 \\ &= \|\dot{p}(t)\|_P^2 - (w_{im}(p(t)), p(t) - p^\dagger)_P, \end{aligned}$$

which implies that (by noting $\dot{\mathcal{E}}(t) = -\eta \|\dot{p}(t)\|_P^2$ and the inequality (B.2))

$$3\dot{\mathcal{E}}(t) + 2\eta\mathcal{E}(t) + \eta\dot{h}(t) = \eta [2V(p(t)) - (p - p^\dagger, w_{im}(p(t)))_P] \leq 0.$$

Integrate the above inequality on $[0, T]$ to obtain together with the non-negativity of $\mathcal{E}(t)$

$$\int_0^T \mathcal{E}(t) dt \leq \frac{3}{2\eta} (\mathcal{E}(0) - \mathcal{E}(T)) - \frac{1}{2} (h(T) - h(0)) \leq \left(\frac{3}{2\eta} \mathcal{E}(0) + \frac{1}{2} h(0) \right) - \frac{1}{2} h(T). \quad (\text{B.9})$$

On the other hand, since both $p(t)$ and $\dot{p}(t)$ are uniform bounded, a constant M exists such that $|h(t)| \leq M$. Hence, letting $T \rightarrow \infty$ in (B.9), we obtain

$$\int_0^\infty \mathcal{E}(t) dt < \infty. \quad (\text{B.10})$$

Hence $\lim_{t \rightarrow \infty} \mathcal{E}(t) = 0$, and, consequently, $\lim_{t \rightarrow \infty} \dot{p}(t) = 0$.

Since $\mathcal{E}(t)$ is non-increasing, we deduce that

$$\int_{T/2}^T \mathcal{E}(t) dt \geq \frac{T}{2} \mathcal{E}(T). \quad (\text{B.11})$$

Using (B.10), the left side of (B.11) tends to 0 when $T \rightarrow \infty$, which implies that $\lim_{T \rightarrow \infty} T\mathcal{E}(T) = 0$. Hence, we conclude $\lim_{T \rightarrow \infty} TV(p(T)) = 0$, which yields the desired result in (iv).

Finally, let us show the long-term behavior of $\ddot{p}(t)$. Integrating the inequality (B.4) from 0 to T we obtain that there exists a real number C' such that for every $t \in [0, \infty)$

$$\dot{\mathcal{E}}(T) + \eta e(T) + \eta \|\dot{p}(T)\|_P^2 + (\eta^2 - 1) \int_0^T \|\dot{p}(t)\|_P^2 dt + \eta \int_0^T \|\ddot{p}(t)\|_P^2 dt \leq C'. \quad (\text{B.12})$$

Since both $e(\cdot)$ and $\dot{e}(\cdot)$ are global bounded (note that $p(t), \dot{p}(t) \in L^\infty([0, \infty), P)$), inequality (B.12) gives $\ddot{p}(t) \in L^2([0, \infty), P)$. The relations $\ddot{p}(t) \in L^\infty([0, \infty), P)$ and $\ddot{p}(t) \rightarrow 0$ as $t \rightarrow \infty$ are obvious by noting assertions (i), (ii), (iv) and the connection equation (9).

Appendix C. Proof of Proposition 3

The case with the damping parameter $\eta(t) = r/t$ can be performed along the lines and using the tools of the proof of Lemma 9. Hence, it suffices to show the case with the fixed damping parameter $\eta(t) = \eta$.

Denote by $p^\delta(t) = p^\delta(x, t)$, and define the Lyapunov function of the differential equation (15) by $\mathcal{E}(t) = V(p^\delta(t)) + \frac{1}{2} \|\dot{p}^\delta(t)\|_P^2$. Similar to the proof of Lemma 5, we have

$$\dot{\mathcal{E}}(t) = -\eta \|\dot{p}^\delta(t)\|_P^2. \quad (\text{C.1})$$

Hence, $\mathcal{E}(t)$ is non-increasing, and consequently, $\|\dot{p}^\delta(t)\|_P^2 \leq 2\mathcal{E}(0)$. Therefore, $\dot{p}^\delta(\cdot)$ is uniform bounded. Integrating both sides in (C.1), we obtain

$$\int_0^\infty \|\dot{p}^\delta(t)\|_P^2 dt \leq \mathcal{E}(0)/\eta < \infty,$$

which yields $\dot{p}^\delta(\cdot) \in L^2([0, \infty), P)$.

Now, let us show that for any $p^\dagger \in P$ the following inequality holds.

$$\limsup_{t \rightarrow \infty} V(p^\delta(t)) \leq V(p^\dagger). \quad (\text{C.2})$$

Consider for every $t \in [0, \infty)$ the function $e(t) = e(t; p^\dagger) := \frac{1}{2} \|p^\delta(t) - p^\dagger\|_P^2$. Since $\dot{e}(t) = (p^\delta(t) - p^\dagger, \dot{p}^\delta(t))_P$ and $\ddot{e}(t) = \|\dot{p}^\delta(t)\|_P^2 + (p^\delta(t) - p^\dagger, \ddot{p}^\delta(t))_P$ for every $t \in [0, \infty)$. Taking into account (15), we get

$$\ddot{e}(t) + \eta \dot{e}(t) + (p^\delta(t) - p^\dagger, w_{im}(p^\delta(t)))_P = \|\dot{p}^\delta(t)\|_P^2. \quad (\text{C.3})$$

On the other hand, by the convexity inequality of the residual norm square functional $V(p^\delta(t))$, we derive

$$V(p^\delta(t)) + (p^\dagger - p^\delta(t), \nabla V(p^\delta(t)))_P \leq V(p^\dagger). \quad (\text{C.4})$$

Combine (C.3) and (C.4) with the definition of $\mathcal{E}(t)$ to obtain

$$\ddot{e}(t) + \eta \dot{e}(t) \leq V(p^\dagger) - \mathcal{E}(t) + \frac{3}{2} \|\dot{p}^\delta(t)\|_P^2.$$

By (C.1), $\mathcal{E}(t)$ is non-increasing, hence, given $t > 0$, for all $\tau \in [0, t]$ we have

$$\ddot{e}(\tau) + \eta \dot{e}(\tau) \leq V(p^\dagger) - \mathcal{E}(t) + \frac{3}{2} \|\dot{p}^\delta(\tau)\|_P^2.$$

By multiplying this inequality with $e^{\eta\tau}$ and then integrating from 0 to θ , we obtain

$$\dot{e}(\theta) \leq e^{-\eta\theta} \dot{e}(0) + \frac{1 - e^{-\eta\theta}}{\eta} (V(p^\dagger) - \mathcal{E}(t)) + \frac{3}{2} \int_0^\theta e^{-\eta(\theta-\tau)} \|\dot{p}^\delta(\tau)\|_P^2 d\tau.$$

Integrate the above inequality once more from 0 to t together with the fact that $\mathcal{E}(t)$ decreases, to obtain

$$e(t) \leq e(0) + \frac{1 - e^{-\eta t}}{\eta} \dot{e}(0) + \frac{\eta t - 1 + e^{-\eta t}}{\eta^2} (V(p^\dagger) - \mathcal{E}(t)) + h(t), \quad (\text{C.5})$$

where $h(t) := \frac{3}{2} \int_0^t \int_0^\theta e^{-\eta(\theta-\tau)} \|\dot{p}^\delta(\tau)\|_P^2 d\tau d\theta$.

Since $e(t) \geq 0$ and $\mathcal{E}(t) \geq V(p^\delta(t))$, it follows from (C.5) that

$$\frac{\eta t - 1 + e^{-\eta t}}{\eta^2} V(p^\delta(t)) \leq e(0) + \frac{1 - e^{-\eta t}}{\eta} \dot{e}(0) + \frac{\eta t - 1 + e^{-\eta t}}{\eta^2} V(p^\dagger) + h(t).$$

Dividing the above inequality by $\frac{\eta t - 1 + e^{-\eta t}}{\eta^2}$ and letting $t \rightarrow \infty$, we deduce that

$$\limsup_{t \rightarrow \infty} V(p^\delta(t)) \leq V(p^\dagger) + \limsup_{t \rightarrow \infty} \frac{\eta}{t} h(t).$$

Hence, for proving (C.2), it suffices to show that $h(\cdot) \in L^\infty([0, \infty), \mathcal{X})$. It is obviously held by noting the following inequalities

$$0 \leq h(t) = \frac{3}{2\eta} \int_0^t (1 - e^{-\eta(t-\tau)}) \|\dot{p}^\delta(\tau)\|_P^2 d\tau \leq \frac{3}{2\eta} \int_0^\infty \|\dot{p}^\delta(\tau)\|_P^2 d\tau < \infty.$$

From the inequality $V(p^\delta(t)) \geq \inf_{p^\dagger \in P} V(p^\dagger)$, we conclude together with (C.2) that

$$\lim_{t \rightarrow \infty} V(p^\delta(t)) = \inf_{p^\dagger \in P} V(p^\dagger). \quad (\text{C.6})$$

Consequently, we have

$$\lim_{t \rightarrow \infty} V(p^\delta(t)) \leq V(p^\dagger) \leq C_0^2 \delta^2.$$

References

- [1] G. Vainikko, A. Veretennikov, *Iteration Procedures in Ill-Posed Problems*, Nauka, 1986, (In Russian).
- [2] U. Tautenhahn, On the asymptotical regularization of nonlinear ill-posed problems, *Inverse Problems* 10 (1994) 1405–1418.
- [3] Y. Zhang, B. Hofmann, On the second order asymptotical regularization of linear ill-posed inverse problems, *Appl. Anal.* (2018) <http://dx.doi.org/10.1080/00036811.2018.1517412>.
- [4] V. Isakov, *Inverse Source Problems*, American Mathematical Society, New York, 1990.
- [5] X. Wang, Y. Guo, D. Zhang, H. Liu, Fourier method for recovering acoustic sources from multi-frequency far-field data, *Inverse Problems* 33 (2017) 035001.
- [6] D. Zhang, Y. Guo, J. Li, H. Liu, Retrieval of acoustic sources from multi-frequency phaseless data, *Inverse Problems* 34 (2018) 094001.
- [7] D. Zhang, Y. Guo, J. Li, H. Liu, Locating multiple multipolar acoustic sources using the direct sampling method, *Commun. Comput. Phys.* 25 (2019) 1328–1356.
- [8] H. Y. Deng, G. Uhlmann, On an inverse boundary problem arising in brain imaging, *J. Differential Equations* 267 (2019) 2471–2502.
- [9] W. Han, W. Cong, G. Wang, Mathematical theory and numerical analysis of bioluminescence tomography, *Inverse Problems* 22 (2006) 1659–1675.
- [10] L. Afraites, M. Dambrine, D. Kateb, Conformal mappings and shape derivatives for the transmission problem with a single measurement, *Numer. Funct. Anal. Optim.* 28 (2007) 519–551.
- [11] S. Song, J. Huang, Solving an inverse problem from bioluminescence tomography by minimizing an energy-like functional, *J. Comput. Anal. Appl.* 14 (2012) 544–558.
- [12] X. Cheng, R. Gong, W. Han, X. Zheng, A novel coupled complex boundary method for inverse source problems, *Inverse Problems* 30 (2014) 055002.
- [13] C. Alves, N. Martins, N. Roberty, Full identification of acoustic sources with multiple frequencies and boundary measurements, *Inverse Prob. Imaging* 3 (2009) 275–294.
- [14] Y. Zhang, R. Gong, M. Gulliksson, X. Cheng, A coupled complex boundary expanding compacts method for inverse source problems, *J. Inverse Ill-Posed Prob.* 27 (2019) 67–86.

- [15] H. Engl, M. Hanke, A. Neubauer, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.
- [16] B. Kaltenbacher, A. Neubauer, O. Scherzer, *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*, Walter de Gruyter GmbH & Co., KG, Berlin, 2008.
- [17] A. Neubauer, On Nesterov acceleration for Landweber iteration of linear ill-posed problems, *J. Inverse Ill-Posed Prob.* 25 (2017) 381–390.
- [18] S. Hubmer, R. Ramlau, Convergence analysis of a two-point gradient method for nonlinear ill-posed problems, *Inverse Problems* 33 (2017) 095004.
- [19] R. Gong, B. Hofmann, Y. Zhang, A new class of accelerated regularization methods, with application to bioluminescence tomography, *Inverse Problems* (2020) <http://dx.doi.org/10.1088/1361-6420/ab730b>.
- [20] W. Su, S. Boyd, E. Candes, A differential equation for modeling Nesterov's accelerated gradient method: Theory and insights, *J. Mach. Learn. Res.* 17 (2016) 1–43.
- [21] R. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.
- [22] Y. Zhang, R. Gong, X. Cheng, M. Gulliksson, A dynamical regularization algorithm for solving inverse source problems of elliptic partial differential equations, *Inverse Problems* 34 (2018) 065001.
- [23] H. Attouch, X. Goudou, P. Redont, The heavy ball with friction method, I. the continuous dynamical system, *Commun. Contemp. Math.* 2 (2000) 1–34.
- [24] Z. Opial, Weak convergence of the sequence of successive approximations for nonexpansive mappings, *Bull. Amer. Math. Soc.* 73 (1967) 591–597.
- [25] H. Attouch, J. Peypouquet, The rate of convergence of Nesterov's accelerated forward-backward method is actually faster than $\mathcal{O}(1/k^2)$, *SIAM J. Optim.* 26 (2016) 1824–1834.
- [26] H. Attouch, Z. Chbani, J. Peypouquet, P. Redont, Fast convergence of inertial dynamics and algorithms with asymptotic vanishing viscosity, *Math. Program.* 168 (2018) 123–175.
- [27] M. Motron, Around the best constants for the Sobolev trace map from $w^{1,2}(\omega)$ into $L^1(\partial\omega)$, *Asymptot. Anal.* 29 (2002) 69–90.
- [28] J. Fernandez Bonder, J. Rossi, Existence results for the p -Laplacian with nonlinear boundary conditions, *J. Math. Anal. Appl.* 263 (2001) 195–223.
- [29] P. Tolksdorf, Regularity for a more general class of quasilinear elliptic equations, *J. Differential Equations* 12 (1984) 126–150.
- [30] J. Vazquez, A strong maximum principle for some quasilinear elliptic equations, *Appl. Math. Optim.* 12 (1984) 191–202.
- [31] M. Hanke, A. Neubauer, O. Scherzer, A convergence analysis of the Landweber iteration for nonlinear ill-posed problems, *Numer. Math.* 72 (1995) 21–37.
- [32] C. Johnson, *Numerical Solution of Partial Differential Equations by the Finite Element Method*, Dover, Mineola, 2009.
- [33] K. Atkinson, W. Han, *Theoretical Numerical Analysis: A Functional Analysis Framework*, third ed., Springer-Verlag, New York, 2009.
- [34] E. Hairer, G. Wanner, C. Lubich, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, second ed., Springer, New York, 2006.