



A novel iterative integration regularization method for ill-posed inverse problems

Ce Huang¹ · Li Wang¹ · Minghui Fu¹ · Zhong-Rong Lu¹ · Yanmao Chen¹

Received: 15 September 2019 / Accepted: 22 December 2019 / Published online: 8 January 2020
© Springer-Verlag London Ltd., part of Springer Nature 2020

Abstract

This paper proposes a new iterative integration regularization method for robust solution of ill-posed inverse problems. The proposed method is motivated from the fact that inversion of a positive definite matrix can be expressed in an integral form. Then, the development of the proposed method is mainly twofold. Firstly, two ways—including the linear iteration and the exponential (2^j) iteration—are invoked to compute the integral, of which the exponential iteration is often preferred due to its fast convergence. Secondly, after stability analysis, the proposed method is shown able to filter out the undesired effect of relatively small singular values, while preserving the desired terms of relatively large singular values, i.e., the proposed method has the guaranteed regularization effect. Numerical examples on three typical ill-posed problems are conducted with detailed comparison to some usual direct and iterative regularization methods. Final results have highlighted the proposed method: (a) due to the iterative nature, the proposed method often turns out to be more efficient than the conventional direct regularization methods including the Tikhonov regularization and the truncated singular value decomposition (TSVD), (b) the proposed method converges much faster than the Landweber method and (c) the regularization effect is guaranteed in the proposed method, while may not be in the conjugate gradient method for least squares problem (CGLS).

Keywords Ill-posed problem · Iterative integration regularization · Filter function · Direct and iterative regularization methods · Force reconstruction · Cauchy problem

1 Introduction

Many engineering and scientific problems require the determination of the unknown input and/or the unknown system parameters of a prescribed system from the known, but generally contaminated, output. To mention a few, inverse reconstruction of the external excitations such as the unsteady aerodynamic forces [1], the ocean wave loads [2] etc from the measured dynamic responses [3, 4] is critical to ensuring the reliability and safety of engineering structures; finite element model updating [5–7] that aims to correct the finite element model parameters using vibration measurements is a prerequisite for finite element analysis of real structures; the classic Cauchy problem [8, 9] that is devoted to reconstruct the field in a certain domain from the

measured data on a part of the boundary typically arises in many physical applications, e.g., acoustic cavity detection, steady-state inverse heat conduction and electro-cardiology. Such problems are referred as inverse problems and are often ill-posed, that is, small contamination errors in the measurement would give rise to large perturbations (or errors) in the solution. For a more comprehensive survey on the ill-posed and inverse problems, refer to Ref. [10].

A great deal of researches have been long focused on robust and efficient solution of the ill-posed problems and a common way to circumvent the ill-posedness is to introduce additional priori knowledge on the solution; this leads to the so-called regularization methods [11]. As Benning and Burger stated at the beginning of their work [11], regularization methods are used to introduce priori knowledge and allow a robust approximation of ill-posed (pseudo-)inverses. In principle, different priori knowledge on the solution corresponds to different regularization methods and according to the classification of priori knowledge, three widely recognized classes of regularization methods have been developed in the literature,

✉ Li Wang
wangli75@mail.sysu.edu.cn

¹ Department of Applied Mechanics and Engineering,
Sun Yat-sen University, Guangzhou 510006,
People's Republic of China

- the continuity/smoothing and bounding properties of the solution are naturally implicated in the ℓ^2 regularization, e.g., the Tikhonov regularization [11–13].
- the sparsity (or specifically the number of non-zeros) constraint of the solution is easily enforced by the ℓ^1 regularization (or called the sparse regularization/compressed sensing) [14–16].
- the edge-preserving and scale-dependent properties of a field solution shall be fulfilled by the total variation regularization [17].

In this paper, the class of ℓ^2 regularization methods is followed up for further investigations. Such class of regularization methods tries to filter out the influence of the noise by introducing singular-value-dependent filter factors that keep approaching 1 for relatively large singular values and tend to 0 for relatively small singular values [18]. Regularization methods in this class are further categorized into [19]:

- *Direct regularization methods* Well-known methods are the truncated singular value decomposition (TSVD) [20] and the Tikhonov regularization [12]. The TSVD seems simple in the analysis; however, due to the requirement for singular value decomposition (SVD) of the matrix, the computation cost involved for the TSVD is often high especially for large-scale problems. On the other hand, the Tikhonov regularization additionally introduces the ℓ^2 penalty term as regularization and then, the solutions under a number of candidate regularization parameters should be obtained, each of which requires solving a linear algebraic equation. The Tikhonov regularization is often more efficient than the TSVD; however, efforts should be made to determine a proper set of candidate regularization parameters.
- *Iterative regularization methods* For large-scale problems, the above direct regularization methods often involve the SVD algorithm or solving a number of algebraic equations and as a result, the computation often becomes prohibitive; this essentially calls for more efficient iterative regularization methods [21, 22]. Among the iterative regularization methods, the Landweber iteration [23] is the classical stationary method. The Landweber iteration is effective because its filter factors resemble those for the Tikhonov regularization; however, it often admits slow convergence [21]. To get faster convergence, the conjugate gradient iteration for least squares problems (CGLS) [19] is naturally invoked. The CGLS is a nonstationary method because its iteration matrix depends not only on the iteration number but also on the right-hand side and there is no analytical expressions of the filter factors. Fortunately, it was shown [21, 24] that under some mild conditions, the CGLS is indeed a regularization method. Nevertheless, the mild conditions may

not be guaranteed and consequently, the CGLS may not work well in some cases [25].

A common feature of the regularization methods is that their performance depends on some regularization parameters, e.g., the truncation order of the TSVD, the Tikhonov regularization parameter and the iteration number of iterative regularization methods. By this way, the key issue in connection with the regularization methods is to find a proper regularization parameter that balances the trade-off between filtering out enough noise (or eliminating the terms corresponding to relatively small singular values) and not losing too much information (or keeping the terms corresponding to relatively large singular values) in the solution [18]. To this end, several regularization selection rules including the discrepancy principle [26], generalized cross-validation (GCV) [27] and the L-curve rule [18] have been available and well recognized. Among them, the discrepancy principle seems to be the simplest rule where the regularization parameter is selected upon the prescribed upper bound of the residual norm; however, the upper bound is often unknown and the eventual solution may be over-smoothed [13]; the GCV has been shown as a promising rule in many applications (see Ref. [28] for instance), but may not give satisfactory results if the errors in the right-hand side are highly correlated; the L-curve rule that locates the regularization parameter at the log–log scale L-corner of the residual norm versus the solution norm curve is, nowadays, a widely used rule due to its theoretically guaranteed performance [13, 18] in the class of ℓ^2 regularization methods and its easy generalization to deal with other classes of regularization methods, e.g., the ℓ^1 regularization [29].

In this paper, a novel iterative integration regularization method is proposed for robust solution of ill-posed problems with the ℓ^2 regularization. The key idea behind is that inversion of a positive definite matrix can be alternatively obtained by integration over a exponential matrix function. Along this line, an iterative procedure is developed to compute the integral and further analysis reveals that such iterative integration procedure can indeed pose the regularization/filtering effect, indicating that the proposed iterative integration is a regularization method. The proposed method is expected to have the following features:

- It belongs to the iterative regularization methods and therefore, shall be efficient for large-scale problems.
- Unlike the conventional CGLS, the filter factors of the proposed method can be explicitly derived and are of the same trend to those of the Tikhonov regularization, that is, the filter factors, lying in the interval [0,1] and increasing monotonically with the singular values, tend to 1 for relatively large singular values and decay to 0 as the singular values approach 0. This to some extent guar-

antees the effectiveness and performance of the proposed method.

- It is shown in the later analysis that the proposed method can degenerate to the Landweber method [23] under the linear expansion of the exponential matrix function and the linear iteration. Notwithstanding, using the exponential (2^j) iteration would accelerate the convergence speed. In this way, the Landweber method can be seen as a particular case of the proposed method, nevertheless, the proposed method can possess faster convergence.

As regards the selection of the regularization parameter (or the number of iteration) for the proposed method, the well-recognized L-curve rule is adopted. As is noteworthy, the proposed method is designated for linear ill-posed problems, of which the known right-hand side is a linear function of the unknowns; nevertheless, there is no constraint to extend the proposed method to nonlinear ill-posed problems because a nonlinear ill-posed problem is often solved iteratively by proper linearization and then, in each iteration, a linear ill-posed problem arises which can be tackled by the proposed method. The similar idea on how to solve the nonlinear ill-posed problems by regularization methods can be found in Ref. [30].

The remainder of this paper is organized as follows. In Sect. 2, the filter function is introduced to describe the behavior of ℓ^2 regularization methods and then, some usual direct and iterative regularization methods are revisited from the perspective of the filter function. In Sect. 3, a new iterative integration regularization method is developed upon an integral form for inversion of a positive definite matrix. In doing so, two ways—the linear iteration and the exponential (2^j) iteration—are invoked to iteratively compute the integral and elaborate analysis on the stability and convergence of the proposed method is presented. It is eventually shown that the proposed method has the guaranteed regularization effect. Section 4 accounts for some typical ill-posed problems, by which the performance of the proposed method as well as the comparison to some usual regularization methods is specifically investigated. Final conclusions are drawn in Sect. 5.

2 Ill-posed problem and regularization

2.1 Ill-posed problem with ℓ^2 regularization

Many inverse problems involve solution of the discrete linear equation

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \quad (1)$$

where the system matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ is of full column rank, i.e., $m \geq n$ and \mathbf{A} is injective, $\mathbf{b} \in \mathbb{R}^m$ is the known output vector and $\mathbf{x} \in \mathbb{R}^n$ is to be determined. Linear equation (1) is said to be ill-posed if the singular values of \mathbf{A} cluster at zero giving rise to huge condition numbers [31]. Note that the output vector is always contaminated with noise, that is,

$$\mathbf{b} = \mathbf{b}_{\text{true}} + \mathbf{e}, \mathbf{b}_{\text{true}} = \mathbf{A}\mathbf{x}_{\text{true}}, \quad (2)$$

where \mathbf{x}_{true} is the true solution and \mathbf{e} denotes noise. Under this circumstance, large errors into the solution \mathbf{x} would be induced by the noise \mathbf{e} for the ill-posed problem (1).

A usual way to circumvent the ill-posedness is to introduce priori knowledge on the smoothness/continuity of the solution, which involves addition of the ℓ^2 penalty term $\|\mathbf{Lx}\|_2^2$ and then, the ℓ^2 regularized formulation can be established for the ill-posed problem (1),

$$\mathbf{x}_\lambda = \arg \min_{\mathbf{x}} \{\|\mathbf{Ax} - \mathbf{b}\|_2^2 + \lambda^2 \|\mathbf{Lx}\|_2^2\}, \quad (3)$$

where the regularization matrix $\mathbf{L} \in \mathbb{R}^{l \times n}$ is often obtained through discretization of some differential operator, and $\lambda \geq 0$ is a positive scalar called the regularization parameter. In actual, the formulation (3) is just the well-known Tikhonov regularization [12].

Usually, the regularization matrix \mathbf{L} is of full row rank, that is, $l \leq n$ and \mathbf{L} is surjective. For instance, to enforce the H^s -continuity/smoothness of the uniformly discretized one-dimensional solution, where $H^s, s = 0, 1, 2$ denotes the usual Sobolev spaces, the regularization matrix \mathbf{L} is often defined as [32],

$$\begin{aligned} \mathbf{L} &= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \in \mathbb{R}^{n \times n}, \text{ for } s = 0, \\ \mathbf{L} &= \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & 1 \end{pmatrix} \in \mathbb{R}^{(n-1) \times n}, \text{ for } s = 1, \\ \mathbf{L} &= \begin{pmatrix} 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & -2 & 1 \end{pmatrix} \in \mathbb{R}^{(n-2) \times n}, \text{ for } s = 2. \end{aligned} \quad (4)$$

In case when \mathbf{L} is a identity matrix \mathbf{I} , the ℓ^2 regularization (3) is of the standard form; while in other cases, a simple transformation can be invoked to transform (3) into the standard form and the procedure is briefly given as follows. To arrive at the standard form, let $\mathbf{Lx} = \mathbf{y} \in \mathbb{R}^l$ and then, due to the surjectivity of \mathbf{L} , the following would hold

$$\mathbf{x} = \mathbf{L}^+ \mathbf{y} + \mathbf{K}_0 \mathbf{z}, \quad (5)$$

where $\mathbf{L}^+ \in \mathbb{R}^{n \times l}$ is the pseudo-inverse of \mathbf{L} , i.e., $\mathbf{LL}^+ = \mathbf{I}$, $\mathbf{K}_o \in \mathbb{R}^{n \times (n-l)}$ represents the null space of \mathbf{L} verifying $\text{rank}(\mathbf{K}_o) = n-l$, $\mathbf{LK}_o = \mathbf{0}$, and $\mathbf{z} \in \mathbb{R}^{n-l}$ can be arbitrary. The details on how to get the matrices \mathbf{L}^+ , \mathbf{K}_o can be found in Ref. [19]. Then, substitution of Eq. (5) into the regularized formulation (3) yields

$$\min_{\mathbf{y}, \mathbf{z}} \{ ||\mathbf{A}(\mathbf{L}^+\mathbf{y} + \mathbf{K}_o\mathbf{z}) - \mathbf{b}||_2^2 + \lambda^2 ||\mathbf{y}||_2^2 \}. \quad (6)$$

Minimization over \mathbf{z} leads to $\mathbf{z} = [(\mathbf{AK}_o)^T(\mathbf{AK}_o)]^{-1}(\mathbf{AK}_o)^T(\mathbf{b} - \mathbf{AL}^+\mathbf{y})$ and then, the standard form problem is obtained

$$\mathbf{y}_\lambda = \arg \min_{\mathbf{y}} \{ ||\bar{\mathbf{A}}\mathbf{y} - \bar{\mathbf{b}}||_2^2 + \lambda^2 ||\mathbf{y}||_2^2 \}. \quad (7)$$

where $\bar{\mathbf{A}} = \mathbf{QAL}^+$, $\bar{\mathbf{b}} = \mathbf{Qb}$, $\mathbf{Q} = \mathbf{I} - (\mathbf{AK}_o)[(\mathbf{AK}_o)^T(\mathbf{AK}_o)]^{-1}(\mathbf{AK}_o)^T$.

Above all, the regularization problem (3) can always be transformed into the standard form and therefore, without loss of generality and for brevity, the analysis will be restricted onto the standard-form regularization problem from here on,

$$\mathbf{x}_\lambda = \arg \min_{\mathbf{x}} \{ ||\mathbf{Ax} - \mathbf{b}||_2^2 + \lambda^2 ||\mathbf{x}||_2^2 \}. \quad (8)$$

2.2 Regularization with filter function

To proceed further, consider the singular value decomposition (SVD) of the matrix \mathbf{A} ,

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T = \sum_{i=1}^n \sigma_i \mathbf{u}_i \mathbf{v}_i^T, \quad (9)$$

where the left and right singular vectors \mathbf{U} and \mathbf{V} are orthonormal, and the singular values in the diagonal matrix Σ are nonnegative and nonincreasing, i.e., $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n > 0$. Then, direct solution of problem (1) in the least-squares sense results in

$$\begin{aligned} \mathbf{x}_{\text{LSQ}} &= \sum_{i=1}^n \frac{(\mathbf{u}_i^T \mathbf{b})}{\sigma_i} \mathbf{v}_i = \sum_{i=1}^n \left\{ \frac{(\mathbf{u}_i^T \mathbf{b}_{\text{true}})}{\sigma_i} \mathbf{v}_i + \frac{(\mathbf{u}_i^T \mathbf{e})}{\sigma_i} \mathbf{v}_i \right\} \\ &= \mathbf{x}_{\text{true}} + \sum_{i=1}^n \frac{(\mathbf{u}_i^T \mathbf{e})}{\sigma_i} \mathbf{v}_i, \end{aligned} \quad (10)$$

where the notations in Eq. (2) are used. Due to the ill-conditionedness, there exist some singular values $\sigma_n, \sigma_{n-1}, \dots, \sigma_{n-q}$ clustering at zero with q an integer and the contamination error corresponding to these singular values $||\mathbf{x}_{\text{LSQ}} - \mathbf{x}_{\text{true}}||_2 \geq \sqrt{\sum_{i=n-q}^n \frac{(\mathbf{u}_i^T \mathbf{e})^2}{\sigma_i^2}}$ is dramatically magnified and even overwhelms the true solution \mathbf{x}_{true} . Practical regularization methods should, therefore, introduce filter factors $0 \leq f_i \leq 1$ [18] into the regularized solution,

$$\mathbf{x}_{\text{reg}} = \sum_{i=1}^n f_i \frac{(\mathbf{u}_i^T \mathbf{b})}{\sigma_i} \mathbf{v}_i. \quad (11)$$

For ℓ^2 regularization, the filter factors are often of the following forms,

$$f_i = f(\sigma_i^2, \alpha), \alpha \in \mathcal{A}, \quad (12)$$

where $f()$ is called the filter function, α denotes some regularization parameter and \mathcal{A} is the space containing all reasonable regularization parameters. Then, the error of the regularized solution is obtained as

$$\mathbf{x}_{\text{reg}} - \mathbf{x}_{\text{true}} = \sum_{i=1}^n f_i \frac{(\mathbf{u}_i^T \mathbf{e})}{\sigma_i} \mathbf{v}_i + \sum_{i=1}^n (f_i - 1) \frac{(\mathbf{u}_i^T \mathbf{b}_{\text{true}})}{\sigma_i} \mathbf{v}_i. \quad (13)$$

Aiming to minimize the error (13), several conditions for the filter function $f(\sigma^2, \alpha)$ should be invoked,

P1 The portion $\sum_{i=1}^n f_i \frac{(\mathbf{u}_i^T \mathbf{e})}{\sigma_i} \mathbf{v}_i$ of the error (13) is mainly controlled by the information corresponding to relatively small singular values $\sigma_n, \sigma_{n-1}, \dots, \sigma_{n-q}$ and to reduce this portion of the error, $f_i, i \in [n, n-1, \dots, n-q]$ should tend to zero so that f_i/σ_i is also approaching zero. In other words, there should hold,

$$f(0, \alpha) = 0, \lim_{\sigma \rightarrow 0^+} \frac{f(\sigma^2, \alpha)}{\sigma} = 0, \quad \forall \alpha \in \mathcal{A}. \quad (14)$$

P2 As with the other portion of the error $\sum_{i=1}^n (f_i - 1) \frac{(\mathbf{u}_i^T \mathbf{b}_{\text{true}})}{\sigma_i} \mathbf{v}_i$, it can be divided into two parts: the first $\sum_{i=n-q}^n (f_i - 1) \frac{(\mathbf{u}_i^T \mathbf{b}_{\text{true}})}{\sigma_i} \mathbf{v}_i \approx \sum_{i=n-q}^n (-1) \frac{(\mathbf{u}_i^T \mathbf{b}_{\text{true}})}{\sigma_i} \mathbf{v}_i$ and the second $\sum_{i=1}^{n-q-1} (f_i - 1) \frac{(\mathbf{u}_i^T \mathbf{b}_{\text{true}})}{\sigma_i} \mathbf{v}_i$ where the approximation in the first part holds due to the condition **P1**. On the one hand, to reduce the first part of the error, the discrete Picard condition should be satisfied [13], i.e., the coefficients $|\mathbf{u}_i^T \mathbf{b}_{\text{true}}|$ on average decay to zero faster than the singular values σ_i , or $|\mathbf{u}_i^T \mathbf{b}_{\text{true}}|/\sigma_i$ tends to 0 for some $i \in [n, n-1, \dots, n-q]$. On the other hand, reducing the second part of the error requires $f_i \rightarrow 1$ for relatively large singular values. In this way, the filter function should verify,

$$\lim_{\sigma \rightarrow +\infty} f(\sigma^2, \alpha) = 1, \quad \forall \alpha \in \mathcal{A}, \quad (15)$$

or weakly, $f(\sigma_1^2, \alpha) \rightarrow 1$, for some α .

P3 The filter function should be able to degenerate into the original solution (10) for well-conditioned problems. To achieve so, there should exist a critical regularization parameter α_{cr} so that

$$f(\sigma^2, \alpha_{\text{cr}}) = \lim_{\alpha \rightarrow \alpha_{\text{cr}}} f(\sigma^2, \alpha) = 1, \forall 0 < \sigma \leq \sigma_1. \quad (16)$$

- P4** The filter function should verify $f(\sigma^2, \alpha) \in [0, 1]$.
***P5** In particular, if $f(\sigma^2, \alpha)$ is continuous with respect to both σ and α , the regularized solution (11) can be alternatively expressed by

$$\mathbf{x}_{\text{reg}} = f(\mathbf{A}^T \mathbf{A}, \alpha) \cdot (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}. \quad (17)$$

Note that the condition ***P5** is only optional and is not necessary in terms of the regularization performance. In what follows, some common direct and iterative ℓ^2 regularization methods are briefly revisited from the viewpoint of the filter function.

2.3 Revisit of some direct and iterative regularization methods

To begin with, the Tikhonov regularization requires exact solution of Eq. (8), giving rise to

$$\mathbf{x}_{\text{Tik}} = (\mathbf{A}^T \mathbf{A} + \lambda^2 \mathbf{I})^{-1} \mathbf{A}^T \mathbf{b} = \sum_{i=1}^n \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} \frac{(\mathbf{u}_i^T \mathbf{b})}{\sigma_i} \mathbf{v}_i. \quad (18)$$

In this way, the filter function for the Tikhonov regularization is found to be

$$f_{\text{Tik}}(\sigma^2, \lambda) = \frac{\sigma^2}{\sigma^2 + \lambda^2}, \quad (19)$$

which obviously satisfies the four conditions **P1–P5** in Sect. 2.2 with $\mathcal{A} = \{\lambda > 0\}$ and the critical regularization parameter $\lambda_{\text{cr}} = 0$. Another popular regularization method is the truncated singular value decomposition (TSVD), by which the solution is simply obtained as

$$\mathbf{x}_{\text{TSVD}} = \sum_{i=1}^k \frac{(\mathbf{u}_i^T \mathbf{b})}{\sigma_i} \mathbf{v}_i, \quad k \leq n, \quad (20)$$

where the truncation order k is just the regularization parameter and then, the filter function for the TSVD turns out to be

$$f_{\text{TSVD}}(\sigma^2, k) = \begin{cases} 0, & \text{if } \sigma < \sigma_{t(k,n)} \\ 1, & \text{otherwise,} \end{cases} \quad (21)$$

where $t(k, n) = \min([k], n)$ with $[k]$ denoting the integer part of k and the regularization parameter space is $\mathcal{A} = \{k \geq 1\}$. The filter function $f_{\text{TSVD}}(\sigma^2, k)$ for the TSVD clearly satisfies the conditions **P1–P4** with the critical regularization parameter $k_{\text{cr}} = n$, but does not verify the optional condition ***P5**. That is to say, the TSVD solution can not be obtained by simple multiplication or inversion on matrix \mathbf{A} , rather the singular value decomposition (SVD) algorithm must be called.

The above two regularization methods—the Tikhonov regularization and the TSVD—belong to the direct regularization methods. As it is well known, the SVD algorithm [33] which is often computationally expensive is required for the TSVD; while for the Tikhonov regularization, the inversion $(\mathbf{A}^T \mathbf{A} + \lambda^2 \mathbf{I})^{-1} \mathbf{A}^T \mathbf{b}$ should be computed under a number of candidate regularization parameters. In total, the computation cost for the direct regularization methods is found expensive, especially for large-scale problems. This essentially calls for more efficient iterative regularization methods [34].

The Landweber method [23] is the classical stationary iterative regularization method and its procedure is simply described as follows:

- **Landweber method**

- set the initial vector $\mathbf{x}_0 = \mathbf{0}$,
- do the stationary iteration to get the solution

$$\mathbf{x}_k = \mathbf{x}_{k-1} + \tau \mathbf{A}^T (\mathbf{b} - \mathbf{A} \mathbf{x}_{k-1}), \quad k = 1, 2, 3, \dots, \quad (22)$$

where τ is a real positive constant and controls the behavior of the iteration.

To analyze the filter function for the Landweber method, insert the SVD of \mathbf{A} into Eq. (22) and then, the k th iteration is found to be

$$\mathbf{x}_{\text{Land}} = \sum_{i=1}^n (1 - (1 - \tau \sigma_i^2)^k) \frac{(\mathbf{u}_i^T \mathbf{b})}{\sigma_i} \mathbf{v}_i \quad (23)$$

from which the convergence of the Landweber method requires $0 < \tau < 2/\sigma_1^2$. It is easily seen from Eq. (23) that the filter function for the Landweber method is

$$f_{\text{Land}}(\sigma^2, k) = 1 - (1 - \tau \sigma^2)^k, \quad (24)$$

where the regularization parameter as the iteration number verifies $k \in \mathcal{A} = \{k \geq 1\}$. Obviously, this filter function (23) satisfies the five conditions **P1–*P5** with the critical regularization parameter $k_{\text{cr}} = +\infty$. In practice, σ_1 can be well and easily estimated, e.g., through the subspace iteration, and then, a proper τ can be determined. Note that during the Landweber iteration (22), only matrix–vector multiplication is involved and thus, the Landweber method would be very efficient for large-scale problems if the number of iteration k is reasonably finite. However, the Landweber method often converges slowly [21] so that k may be a large number. To further accelerate the convergence, the conjugate gradient (CG) algorithm to solve the least-squares equation $\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{b}$ is naturally invoked, which is designated as the CGLS:

- **Conjugate gradient algorithm for least squares problem (CGLS)**

- Let \mathbf{x}_0 denote the initial vector that often equals to $\mathbf{0}$, and define two auxiliary vectors $\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0$ and $\mathbf{d}_0 = \mathbf{A}^T\mathbf{r}_0$,
- do the CG iteration for $k = 1, 2, \dots$

$$\begin{aligned}\alpha_k &= \frac{\|\mathbf{A}^T\mathbf{r}_{k-1}\|_2^2}{\|\mathbf{A}\mathbf{d}_{k-1}\|_2^2}, \\ \mathbf{x}_k &= \mathbf{x}_{k-1} + \alpha_k \mathbf{d}_{k-1}, \\ \mathbf{r}_k &= \mathbf{r}_{k-1} - \alpha_k \mathbf{A}\mathbf{d}_{k-1}, \\ \beta_k &= \frac{\|\mathbf{A}^T\mathbf{r}_k\|_2^2}{\|\mathbf{A}^T\mathbf{r}_{k-1}\|_2^2}, \\ \mathbf{d}_k &= \mathbf{A}^T\mathbf{r}_k + \beta_k \mathbf{d}_{k-1}.\end{aligned}\quad (25)$$

The CGLS is a non-stationary method and the filter function is of the following form [35],

$$f_{\text{CGLS}}(\sigma^2, k) = 1 - \prod_{j=1}^k \frac{(\theta_j^{(k)})^2 - \sigma^2}{(\theta_j^{(k)})^2}, \quad k \leq n, \quad (26)$$

where $(\theta_j^{(k)})^2$ are the k eigenvalues of $\mathbf{A}^T\mathbf{A}$ restricted to the $\mathcal{K}_k(\mathbf{A}^T\mathbf{A}, \mathbf{A}^T\mathbf{b}) = \text{span}(\mathbf{A}^T\mathbf{b}, (\mathbf{A}^T\mathbf{A})\mathbf{A}^T\mathbf{b}, \dots, (\mathbf{A}^T\mathbf{A})^{k-1}\mathbf{A}^T\mathbf{b})$. The CGLS often possesses faster convergence (or less k) than the Landweber method. As k increases, the eigenvalues converge to some of the eigenvalues of $\mathbf{A}^T\mathbf{A}$, i.e., $\theta_j^{(k)} \approx \sigma_i$ for selected i, j . By Eq. (26), the filter function for the CGLS is also dependent on the right-hand side \mathbf{b} . Turning to the conditions in Sect. 2.2, $f_{\text{CGLS}}(\sigma^2, k)$ obviously verifies conditions **P1**, **P2**, ***P5** with $k_{\text{cr}} = n$ and approximately satisfies condition **P3**, but does not verify the condition **P4** because $f_{\text{CGLS}}(\sigma^2, k)$ may be greater than 1 [25]. As a consequence, the regularization effect of the CGLS is not always guaranteed [25].

In the next section, a new iterative integration regularization method is to be developed. It is iterative and stationary and would, therefore, be efficient and possess the guaranteed regularization effect.

3 The iterative integration regularization method

3.1 Inversion by iterative integration

To start with, an alternative integral formulation for inversion of the matrix $\mathbf{A}^T\mathbf{A}$ is presented in the following theorem.

Theorem 1 Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ be a matrix with full column rank, i.e., $\text{rank}(\mathbf{A}) = n$ and then, an integral form for inversion of $\mathbf{A}^T\mathbf{A}$ holds true,

$$(\mathbf{A}^T\mathbf{A})^{-1} = \int_0^{+\infty} e^{-(\mathbf{A}^T\mathbf{A})t} dt. \quad (27)$$

Proof Considering the full column rank property of \mathbf{A} and the SVD as in equation (9), it is easily obtained that

$$(\mathbf{A}^T\mathbf{A}) = \mathbf{V}\Sigma^2\mathbf{V}^T \quad (28)$$

where \mathbf{V}, Σ are as illustrated below (9). By Eq. (28), one can get

$$\int_0^{+\infty} e^{-(\mathbf{A}^T\mathbf{A})t} dt = \mathbf{V} \int_0^{+\infty} e^{-\Sigma^2 t} dt \mathbf{V}^T = \mathbf{V}\Sigma^{-2}\mathbf{V}^T \quad (29)$$

from which the equality (27) follows trivially.

Now turn to the ill-posed problem (1), and the least-squares solution requires solving $\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$. By Theorem 1, the solution is found to be

$$\mathbf{x} = \int_0^{+\infty} e^{-(\mathbf{A}^T\mathbf{A})t} \mathbf{A}^T \mathbf{b} dt. \quad (30)$$

In this way, the solution shall be approximately obtained by introducing some finite (and often large) number $T_0 > 0$,

$$\mathbf{x}_{\text{int}} = \int_0^{T_0} e^{-(\mathbf{A}^T\mathbf{A})t} \mathbf{A}^T \mathbf{b} dt. \quad (31)$$

and how to practically compute the integral (31) in an iterative manner will be elaborated below.

Let $\tau > 0$ be small enough and then, by Taylor expansion, it is obtained that

$$\begin{aligned}e^{-(\mathbf{A}^T\mathbf{A})\tau} &\approx \mathbf{B}_{p,\tau} := \mathbf{I} - (\mathbf{A}^T\mathbf{A})\tau + \frac{(\mathbf{A}^T\mathbf{A})^2\tau^2}{2!} \\ &\quad - \frac{(\mathbf{A}^T\mathbf{A})^3\tau^3}{3!} + \dots + (-1)^p \frac{(\mathbf{A}^T\mathbf{A})^p\tau^p}{p!}\end{aligned}\quad (32)$$

and then,

$$\begin{aligned}\int_0^\tau e^{-(\mathbf{A}^T\mathbf{A})t} dt &= (\mathbf{A}^T\mathbf{A})^{-1}(\mathbf{I} - e^{-(\mathbf{A}^T\mathbf{A})\tau}) \\ &\approx \mathbf{C}_{p,\tau} := (\mathbf{A}^T\mathbf{A})^{-1}(\mathbf{I} - \mathbf{B}_{p,\tau}) \\ &= \tau\mathbf{I} - \frac{(\mathbf{A}^T\mathbf{A})\tau^2}{2!} + \frac{(\mathbf{A}^T\mathbf{A})^2\tau^3}{3!} \\ &\quad - \frac{(\mathbf{A}^T\mathbf{A})^3\tau^4}{4!} + \dots + (-1)^{p-1} \frac{(\mathbf{A}^T\mathbf{A})^{p-1}\tau^p}{p!},\end{aligned}\quad (33)$$

where p denotes the truncation order of Taylor expansion. To go further, set

$$\mathbf{x}_k = \int_0^{k\tau} e^{-(\mathbf{A}^T \mathbf{A})t} \mathbf{A}^T \mathbf{b} dt \quad (34)$$

and then, two ways can be invoked to get the solution iteratively with the help of the two approximations (32) and (33):

- **Iterative integration-linear(IIL)** At first, a simple equality is obtained,

$$\begin{aligned} \mathbf{x}_k &= (\mathbf{A}^T \mathbf{A})^{-1} (\mathbf{I} - e^{-(\mathbf{A}^T \mathbf{A})k\tau}) \mathbf{A}^T \mathbf{b} \Rightarrow e^{-(\mathbf{A}^T \mathbf{A})k\tau} \mathbf{A}^T \mathbf{b} \\ &= \mathbf{A}^T \mathbf{b} - \mathbf{A}^T \mathbf{A} \mathbf{x}_k. \end{aligned} \quad (35)$$

Then, by the equality (35), the iterative procedure to get the solution (34) can be established with $\mathbf{x}_0 = \mathbf{0}$ and for $k = 1, 2, \dots$

$$\begin{aligned} \mathbf{x}_k &= \mathbf{x}_{k-1} + \int_{(k-1)\tau}^{k\tau} e^{-(\mathbf{A}^T \mathbf{A})t} \mathbf{A}^T \mathbf{b} dt \\ &= \mathbf{x}_{k-1} + \int_0^\tau e^{-(\mathbf{A}^T \mathbf{A})t} dt \cdot e^{-(\mathbf{A}^T \mathbf{A})(k-1)\tau} \mathbf{A}^T \mathbf{b} \\ &= \mathbf{x}_{k-1} + \int_0^\tau e^{-(\mathbf{A}^T \mathbf{A})t} dt \cdot (\mathbf{A}^T \mathbf{b} - \mathbf{A}^T \mathbf{A} \mathbf{x}_{k-1}) \\ &= \mathbf{x}_{k-1} + \mathbf{C}_{p,\tau} \mathbf{A}^T (\mathbf{b} - \mathbf{A} \mathbf{x}_{k-1}), \end{aligned} \quad (36)$$

where the matrix $\mathbf{C}_{p,\tau}$ is defined in Eq. (33), $p \geq 1$ is an integer and τ is a proper parameter; both p and τ govern the performance of the iteration. The IIL is named after the fact that the number $T_0 = k\tau$ in the iteration (36) increases linearly with the number of iterations.

- **Iterative integration-exponential (IIE)** Distinct to the linear iteration in IIL, the number $T_0 = 2^j \tau$ is expected to increase exponentially herein. This follows the 2^j algorithm that has been well used in the precise integration method [36]. To do so, set $\mathbf{T}_j = e^{-(\mathbf{A}^T \mathbf{A})2^j \tau} - \mathbf{I}$ and simple analysis gives rise to

$$\mathbf{x}_{2^j} = (\mathbf{I} + e^{-(\mathbf{A}^T \mathbf{A})2^{j-1}\tau}) \int_0^{2^{j-1}\tau} e^{-(\mathbf{A}^T \mathbf{A})t} dt \mathbf{A}^T \mathbf{b}. \quad (37)$$

Then, the iterative procedure is given as: initially set $\mathbf{T}_0 = \mathbf{B}_{p,\tau} - \mathbf{I}$ and $\mathbf{x}_{2^0} = \mathbf{C}_{p,\tau} \mathbf{A}^T \mathbf{b}$, and then, proceed for $j = 1, 2, \dots$,

$$\begin{aligned} \mathbf{x}_{2^j} &= (2\mathbf{I} + \mathbf{T}_{j-1}) \mathbf{x}_{2^{j-1}}, \\ \mathbf{T}_j &= 2\mathbf{T}_{j-1} + \mathbf{T}_{j-1} \cdot \mathbf{T}_{j-1}. \end{aligned} \quad (38)$$

Note that both the IIL and the IIE are stationary and are derived as approximations of (34), and their performance is dependent on two parameters p, τ . In the following theorem, the condition on p, τ such that the two iterations—the IIL (36) and the IIE (38) give convergent solutions to $(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}$ is presented.

Theorem 2 If $p \geq 1, \tau > 0$ are selected to verify

$$\rho(\mathbf{B}_{p,\tau}) < 1 \quad (39)$$

with $\mathbf{B}_{p,\tau}$ defined in Eq. (32) and $\rho(\cdot)$ denoting the spectral radius of a matrix, the two iterations—the IIL (36) and the IIE (38) would give convergent solutions to $(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}$.

Proof At first, it is easily found from Eq. (33) that

$$\mathbf{B}_{p,\tau} = \mathbf{I} - (\mathbf{A}^T \mathbf{A}) \mathbf{C}_{p,\tau}, \quad (40)$$

where $p \geq 1$ is an integer. Then, further analysis on the IIL (36) yields

$$\begin{aligned} \mathbf{x}_k &= \mathbf{C}_{p,\tau} \mathbf{A}^T \mathbf{b} + \mathbf{B}_{p,\tau} \mathbf{x}_{k-1} \\ &= [\mathbf{I} + \mathbf{B}_{p,\tau} + \cdots + (\mathbf{B}_{p,\tau})^{k-1}] \mathbf{C}_{p,\tau} \mathbf{A}^T \mathbf{b} \\ &= (\mathbf{I} - (\mathbf{B}_{p,\tau})^k) (\mathbf{I} - \mathbf{B}_{p,\tau})^{-1} \mathbf{C}_{p,\tau} \mathbf{A}^T \mathbf{b} \\ &= (\mathbf{I} - (\mathbf{B}_{p,\tau})^k) (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} \end{aligned} \quad (41)$$

and under the condition (39), it is naturally deduced that

$$\lim_{k \rightarrow +\infty} (\mathbf{I} - (\mathbf{B}_{p,\tau})^k) (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}. \quad (42)$$

On the other hand, for the IIE (38), simple manipulations give rise to

$$\begin{aligned} \mathbf{I} + \mathbf{T}_j &= (\mathbf{I} + \mathbf{T}_{j-1})^2 = (\mathbf{I} + \mathbf{T}_0)^{2^j} = (\mathbf{B}_{p,\tau})^{2^j}, \\ \mathbf{x}_{2^j} &= (\mathbf{I} + (\mathbf{B}_{p,\tau})^{2^{j-1}}) \mathbf{x}_{2^{j-1}} = [\prod_{s=0}^{j-1} (\mathbf{I} + (\mathbf{B}_{p,\tau})^{2^s})] \mathbf{x}_{2^0} \\ &= [\mathbf{I} + (\mathbf{B}_{p,\tau}) + (\mathbf{B}_{p,\tau})^2 + \cdots + (\mathbf{B}_{p,\tau})^{2^{j-1}}] \mathbf{C}_{p,\tau} \mathbf{A}^T \mathbf{b} \\ &= (\mathbf{I} - (\mathbf{B}_{p,\tau})^{2^j}) (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} \end{aligned} \quad (43)$$

which certainly tends to $(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}$ as $j \rightarrow +\infty$ under the condition (39).

Remark 1 Paying attention to the IIL (36), if the linear Taylor truncation order is adopted in Eqs. (32) or (33), i.e., $p = 1$, there is $\mathbf{C}_{p,\tau} = \tau \mathbf{I}$ and then, the IIL (36) turns out to be $\mathbf{x}_k = \mathbf{x}_{k-1} + \tau \mathbf{A}^T (\mathbf{b} - \mathbf{A} \mathbf{x}_{k-1})$, which coincides perfectly with the Landweber method (22). Clearly, the proposed iterative integration method degenerates to the Landweber method under the linear Taylor expansion and the linear iteration. In other words, the proposed method can to some extent viewed as the extension of the Landweber method.

Remark 2 Comparing (41) with (43), the final results by the IIL and the IIE are of the same form, except that the ways to do iterations are different: the IIL is the linear (k) iteration, while the IIE is the exponential (2^j) iteration. In a single iteration, the computation cost for the IIL is much less than that for the IIE, because the IIL only involves the matrix–vector multiplication, while the IIE requires the matrix–matrix

multiplication. However, the IIE converges much faster than the IIL especially for the large-scale ill-posed problems. Specifically, if the iteration number k is finite and quite less than n , the IIL would be much efficient than the IIE; while if too many iterations are required ($k > n \ln k$), the IIE is often preferred.

Remark 3 Focusing on the condition (39), it is easily obtained that for $p \geq 0$

$$\begin{aligned}\rho(\mathbf{B}_{p,\tau}) &= \max\{|g_p(\tau\sigma^2)|, \sigma \in \{\sigma_1, \sigma_2, \dots, \sigma_n\}\}; \\ g_p(t) &= 1 - t + \frac{t^2}{2!} + \dots + (-1)^p \frac{t^p}{p!},\end{aligned}\quad (44)$$

where $g_p(t)$ is the p th order Taylor expansion of e^{-t} . To go further, set $r_p(t) = e^{-t} - g_p(t)$ and the following hold

$$r_p(0) = 0, r'_p(t) = -r_{p-1}(t). \quad (45)$$

Then, for $t \geq 0$, one can have $r_0(t) \leq 0$ and then, recursively for $k = 1, 2, \dots$

$$\begin{aligned}r_{2k-2}(t) &\leq 0; \\ r'_{2k-1}(t) &= -r_{2k-2}(t) \geq 0 \Rightarrow r_{2k-1}(t) \geq 0; \\ r'_{2k}(t) &= -r_{2k-1}(t) \leq 0 \Rightarrow r_{2k}(t) \leq 0.\end{aligned}\quad (46)$$

This gives rise to,

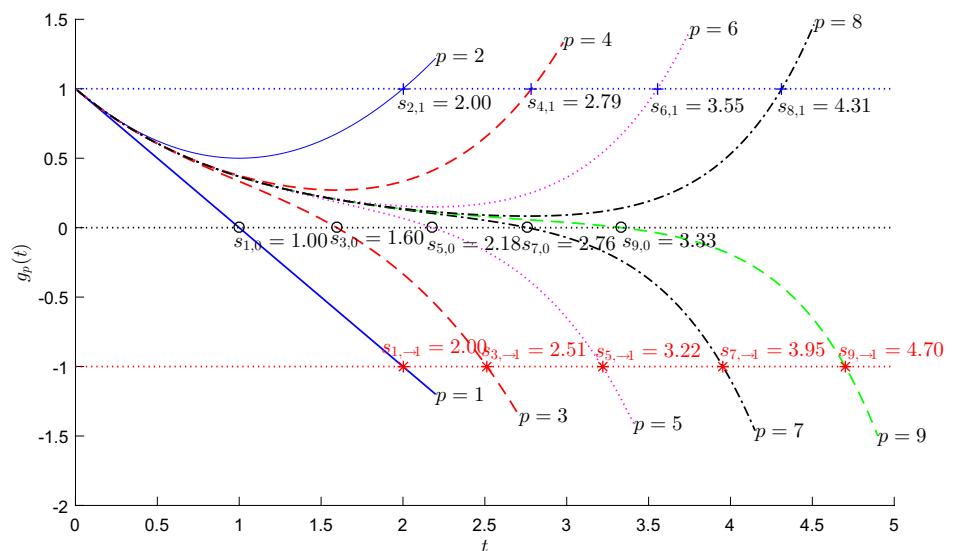
$$g_{2k-1}(t) \leq e^{-t} \leq g_{2k}(t), g_{2k}(t) > 0. \quad (47)$$

Based on the results in Eq. (47), there are

$$g'_{2k+1}(t) = -g_{2k}(t) < 0; g''_{2k+2}(t) = g_{2k}(t) > 0 \quad (48)$$

indicating that for $t \in [0, \tau\sigma_1^2]$, $g_{2k-1}(t)$ is a monotonically decreasing function and $g_{2k}(t)$ is a strictly convex function.

Fig. 1 Schematic of function $g_p(t)$



Along with the fact $g_{2k-1}(0) = g_{2k}(0) = 1$ and Eqs. (44) and (48), the condition (39) turns into

$$|g_p(\tau\sigma_1^2)| < 1 \Leftrightarrow \begin{cases} g_p(\tau\sigma_1^2) > -1, & \text{if } p = 2k - 1, \\ g_p(\tau\sigma_1^2) < 1, & \text{if } p = 2k. \end{cases} \quad (49)$$

For brevity, let the real solution of $g_p(t) = z, t > 0$ under some proper z be denoted by $s_{p,z}$ and for a schematic view of these solutions, refer to Fig. 1. Consequently, the condition (39) is simplified to

$$0 < \tau < \tau_p^{\text{Upper}} := \begin{cases} \frac{s_{p-1}}{\sigma_1^2}, & \text{for } p \text{ odd,} \\ \frac{s_{p,1}}{\sigma_1^2}, & \text{for } p \text{ even,} \end{cases} \quad (50)$$

and the greater p is, the larger upper bound τ_p^{Upper} would be. In this paper, it is set $\tau = 0.8\tau_p^{\text{Upper}}$.

Up till now, the iterative integration method with two variants—the IIL and the IIE—has been developed for inverse solution of the ill-posed problem (1). In what follows, the iterative integration method will be shown to pose the regularization effect.

3.2 Regularization effect

This section starts with the ideal case (31) of the integration method with $T_0 = k\tau$,

$$\mathbf{x}_{\text{int}} = [\mathbf{I} - e^{-(\mathbf{A}^T \mathbf{A})k\tau}] (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} \quad (51)$$

and then, the filter function for the ideal case is found to be

$$f_{\text{int}}(\sigma^2, k) = 1 - e^{-\sigma^2 k \tau}. \quad (52)$$

While turning to equations (41) and (43), the filter functions for the two practical variants of the iterative integration method—the IIL and the IIE are

$$\begin{aligned} f_{\text{IIL}}(\sigma^2, k) &= 1 - (g_p(\tau\sigma^2))^k, \\ f_{\text{IIE}}(\sigma^2, j) &= 1 - (g_p(\tau\sigma^2))^{2^j}, \end{aligned} \quad (53)$$

where the function $g_p(\cdot)$ has been defined in Eq. (44). Clearly, $f_{\text{IIL}}(\sigma^2, k)$ satisfies the five conditions in Sect. 2.2, indicating that the ideal case is indeed a regularization method.

As with the two practical variants—the IIL and the IIE, they have nearly the same filter functions, except for the different exponentials k and 2^j . Thus, without loss of generality, the focus is on the analysis of the filter function $f_{\text{IIL}}(\sigma^2, k)$ of the IIL. At first, considering the fact that $g_p(0) = 1, g'_p(0) = -1$, there is

$$f_{\text{IIL}}(0, k) = 0, \lim_{\sigma \rightarrow 0} \frac{f_{\text{IIL}}(\sigma^2, \alpha)}{\sigma} = \lim_{\sigma \rightarrow 0} \frac{-k(g_p(\tau\sigma^2))^{k-1}g'_p(\tau\sigma^2)2\tau\sigma}{1} = 0 \quad (54)$$

meaning that the condition **P1** is satisfied. Secondly, under the condition (50), there is $|g_p(\tau\sigma^2)| < 1, \forall \sigma \in [\sigma_n, \sigma_1]$ and then, the condition **P2** is verified under some large k . Other conditions **P3**, **P4**, ***P5** are obviously fulfilled by $f_{\text{IIL}}(\sigma^2, \alpha)$ with the critical parameter $k_{\text{cr}} = +\infty$. To conclude, the iterative integration method with two variants has been shown to satisfy the five conditions in Sect. 2.2 and as a result, it is a strict regularization method.

At the closure of this sub-section, the filter functions for the six regularization methods involved in this work—the Tikhonov regularization, the TSVD, the Landweber method, the CGLS, and the proposed IIL and IIE are summarized in Table 1 for comparison. In the next section, how to select the regularization parameter for the iterative integration method is elaborated.

3.3 Selection of regularization parameter using the L-curve

In the above, the iterative integration method has been shown to pose the regularization effect and the iteration number k for the IIL (resp. j for the IIE) is just the regularization parameter, that remains to be determined.

The L-curve has been well recognized as an effective way to select the regularization parameter [18]. To do so, for the IIL, the norm of the regularized solution and the norm of the residual should be calculated at each iteration k ,

$$\rho^2(k) = \|\mathbf{Ax}_k - \mathbf{b}\|_2^2, \eta^2(k) = \|\mathbf{x}_k\|_2^2. \quad (55)$$

Then, plot the log–log scale curve

$$(\log \rho(k), \log \eta(k)) \quad (56)$$

and the optimal k is chosen at the L-corner of the curve. As with how to find the L-corner, the cubic spline curve fitting [18] or the triangle method [37] can be used. The above procedure can work for the IIE by simply replacing \mathbf{x}_k in (55) with \mathbf{x}_{2^j} . Note that the iteration should go a few iterations too far from the optimal choice in order to determine the L-corner of the curve.

4 Numerical examples and applications

In this section, three typical ill-posed problems arising from some real applications are investigated to see the effectiveness and efficiency of the proposed iterative integration regularization method including the IIL and the IIE. Other regularization methods as listed in Table 1 will also be considered for comparison. The root-mean-square (RMS) noise is added to the data as follows

$$\mathbf{b} = \mathbf{b}_{\text{true}} + e \cdot \text{RMS}(\mathbf{b}_{\text{true}}) \cdot \mathbf{r}, \quad (57)$$

where $\text{RMS}(\mathbf{b}) = \sqrt{(b_1^2 + b_2^2 + \dots + b_m^2)/m}$, e is the noise level (e.g., = 1%), and \mathbf{r} is a random vector having the same

Table 1 The filter functions for various regularization methods

Method	Abbreviation	Filter function
Tikhonov	Tik	$f_{\text{Tik}}(\sigma^2, \lambda) = \frac{\sigma^2}{\sigma^2 + \lambda^2}$
Truncated singular value decomposition	TSVD	$f_{\text{TSVD}}(\sigma^2, k) = \begin{cases} 0, & \text{if } \sigma < \sigma_{t(k,n)} \\ 1, & \text{otherwise} \end{cases}$
Landweber method	Land	$f_{\text{Land}}(\sigma^2, k) = 1 - (1 - \tau\sigma^2)^k$
CG algorithm for least squares problem	CGLS	$f_{\text{CGLS}}(\sigma^2, k) = 1 - \prod_{j=1}^k \frac{(\theta_j^{(k)})^2 - \sigma^2}{(\theta_j^{(k)})^2}$
Iterative integration method-Linear	IIL	$f_{\text{IIL}}(\sigma^2, k) = 1 - (g_p(\tau\sigma^2))^k$
Iterative integration method-Exponential	IIE	$f_{\text{IIE}}(\sigma^2, j) = 1 - (g_p(\tau\sigma^2))^{2^j}$

size with b and pertaining to the standard normal distribution.

4.1 Fredholm integral equation of the first kind

Consider the Fredholm integral equation of the first kind [38]:

$$\int_a^b K(s, t)f(t)dt = g(s), c \leq s \leq d, \quad (58)$$

where $K(s, t)$ is the kernel function, $g(s), c \leq s \leq d$ is the known right-hand side and $f(t), a \leq t \leq b$ is the unknown solution remaining to be determined. The Fredholm integral equation is widely encountered in many applications, including image reconstruction [39], inverse Laplace transformation and so on. Usually, this equation is solved through the following discretization: let $a = t_1 < t_2 < \dots < t_n = b, c = s_1 < s_2 < \dots < s_m = d$ be the (usually uniform) discretization of the intervals $[a, b]$ and $[c, d]$. Then, set $\mathbf{x} = [f(t_1); f(t_2); \dots; f(t_n)]$ and $\mathbf{b} = [g(s_1); g(s_2); \dots; g(s_m)]$, and by the Trapezoid integration rule, a discrete equation of the form (1) is naturally invoked for the Fredholm integral Eq. (58) with

$$\begin{aligned} \mathbf{A} &= (a_{ij} := K(s_i, t_j)\Delta t_j)_{m \times n}, \Delta t_j \\ &= \begin{cases} (t_2 - t_1)/2, & j = 1, \\ (t_{j+1} - t_{j-1})/2, & j = 2, \dots, n-1, \\ (t_n - t_{n-1})/2, & j = n. \end{cases} \end{aligned} \quad (59)$$

Herein, the particular case is a one-dimensional problem in image reconstruction [13] where the image is blurred by a known point-spread function. The kernel function is the point-spread function of an infinitely long slit, given by

$$K(s, t) = [(\cos s + \cos t) \cdot \text{sinc}(\sin s + \sin t)]^2, -\pi/2 \leq s \leq \pi/2 \quad (60)$$

with $\text{sinc}(x) = \frac{\sin \pi x}{\pi x}$, while the desired solution is set to

$$f(t) = \exp(-4(t - 0.5)^2) + \exp(-4(t + 0.5)^2), -\pi/2 \leq t \leq \pi/2. \quad (61)$$

For discretization, $m = n = 64$ equally spaced points in $[-\pi/2, \pi/2]$ are used for both t and s and then, the system matrix \mathbf{A} is obtained as in Eq. (59). To see more into the system matrix \mathbf{A} , the singular values of \mathbf{A} are depicted in Fig. 2. As it is observed, there are many singular values being less than 1×10^{-15} , that is, clustering at zero. The condition number as the ratio of the largest singular value to the smallest $\sigma_1/\sigma_n = 4.6 \times 10^{16}$ is huge, indicating that the problem is indeed ill-posed. In this example, the true solution \mathbf{x}_{true} can be obtained based on (61) and then, the right-hand side is obtained by adding noise with level $e = 1\%$ to the true one $\mathbf{b}_{\text{true}} = \mathbf{A}\mathbf{x}_{\text{true}}$ in the form of (57). For simplicity, it is taken

$\mathbf{L} = \mathbf{I}$, the identity matrix in the regularization formulation (3).

The six regularization methods abbreviated as listed in Table 1 are adopted to solve the ill-posed problem. Note that the Taylor truncation order for the proposed two variants of the iterative integration method—the IIL and the IIE, is fixed at $p = 3$ (refer to Eq. (32)) at first. The L-curve rule is used to select a proper regularization parameter and the detailed L-curves for the six regularization methods are displayed in Fig. 3 where the selected regularization parameters at the L-corner are highlighted in red. As it is seen, among the iterative regularization methods, both the Landweber method and the IIL ($p = 3$) require over 2500 iterations to reach the L-corner, while the CGLS and the IIE ($p = 3$) arrives at the L-corner within only 12 iterations. Note that the Landweber method is a particular case of the IIL but with Taylor truncation order being $p = 1$ and the slow convergence of the IIL for this ill-posed problem is observed. In contrast, fast convergence is seen by the CGLS and the IIE ($p = 3$). Except for the Landweber method and the IIL ($p = 3$), the L-shape of the residual norm vs. the solution norm curve for other regularization methods is quite clear and the selection at the L-corner seems straightforward.

With the regularization parameters selected at the L-corner, the solutions by the six regularization methods are obtained as graphically shown in Fig. 4. Clearly, all the six regularization methods have led to satisfactory solution of the Fredholm integral equation. To further quantify the performance of the six regularization methods for the Fredholm integral equation, the relative error (RE) is calculated as

$$\text{RE} = \frac{\|\mathbf{x}_{\text{reg}} - \mathbf{x}_{\text{true}}\|_2}{\|\mathbf{x}_{\text{true}}\|_2} \times 100\%, \quad (62)$$

with \mathbf{x}_{reg} denoting the solution by some regularization method, and the computation time is also counted; detailed results are presented in Table 2. The relative errors for the

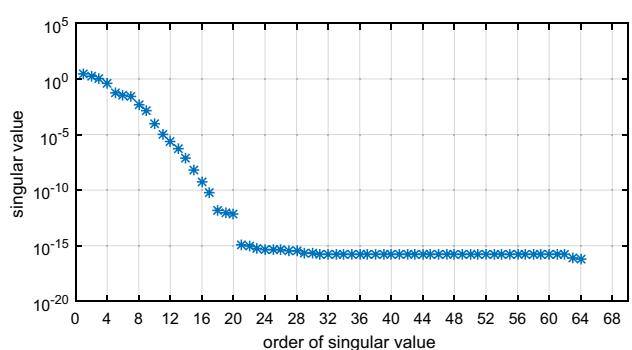


Fig. 2 Schematic of the singular values of the system matrix for Fredholm integral equation problem

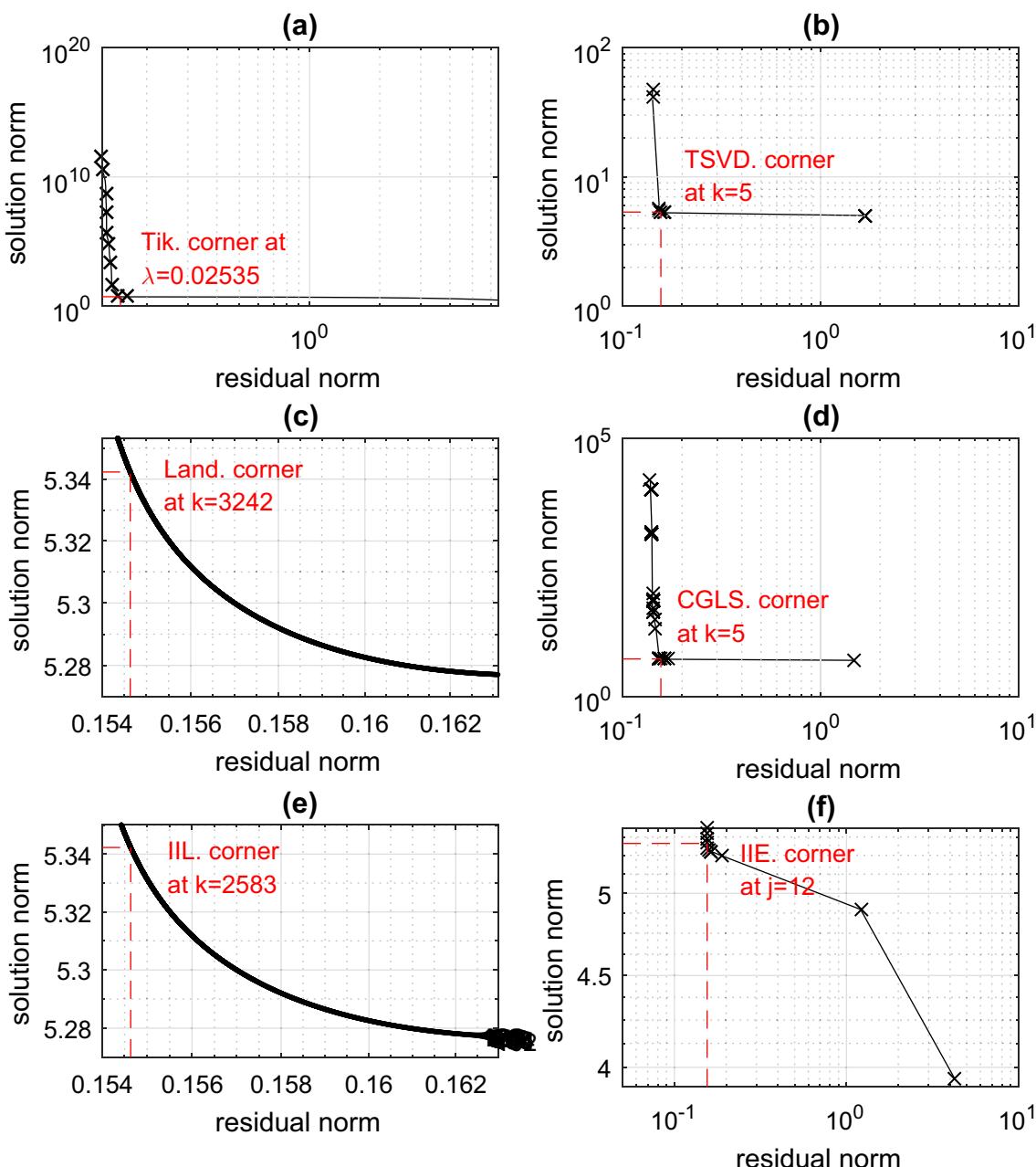


Fig. 3 L-curves of different regularization methods for Fredholm integral equation problem: **a** Tik, **b** TSVD, **c** Land, **d** CGLS, **e** IIL ($p = 3$) and **f** IIE ($p = 3$)

six regularization methods approaching to each other are all less than 8%, meaning that this Fredholm integral equation is well solved. Among them, the proposed IIE ($p = 3$) gives rise to least relative error (RE = 5.88%); the effectiveness of the proposed iterative integration regularization method (including the IIE and the IIL) for the ill-posed Fredholm integral equation is verified. As with the computation time, the cost for the six regularization methods is at the same

cheap level due to the fact the system matrix \mathbf{A} is only of small size $m = n = 64$, nonetheless, the iterative methods including the CGLS and the IIE turn out to be slightly more efficient. To further visual the regularization effect of the six regularization methods, the filter factors (see Table 1) under different singular values are plotted in Fig. 5. Except for the CGLS, other five methods share the same trend of the filter factors: the filter factors tend to zero for relatively

small singular values ($\sigma_i \leq 1 \times 10^{-3}$) to filter out undesired noise effect, the filter factors tend to one for relatively large singular values ($\sigma_j > 0.1$) to keep the terms of the true solution and the filter factors lie almost monotonically between [0, 1]. Particular attention should be paid to the filter factors of the CGLS where the filter factor reaches 1.8 at the 4th singular value, being undesirably greater than 1 and this indicates that the CGLS is not a strict regularization method. Despite of the undesirable trend in the filter function of the CGLS, the CGLS still performs well for this Fredholm

integral equation and the reason may reside at that the term corresponding to the 4th singular value is negligible in the true solution.

To conclude, for the proposed iterative integration regularization method, the IIL admits slow convergence, while fast convergence is often reached by the IIE. In this way, the exponential iteration (IIE) is often preferred to carry out the proposed integration regularization. The IIE is found to give the solution with least relative error, and due to iterative nature, the IIE and the CGLS are slightly more efficient in this example than the direct regularization

Fig. 4 Plot of the solutions by the six regularization methods for Fredholm integral equation problem

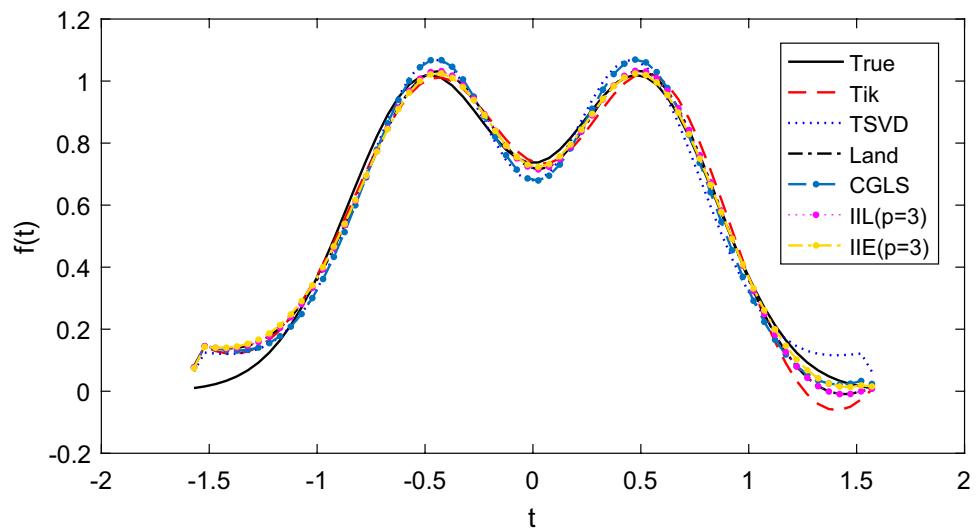
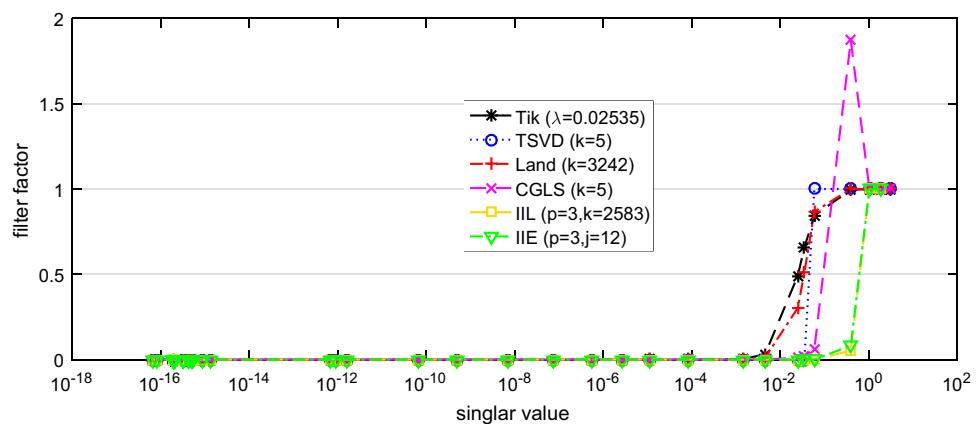


Table 2 Results on relative error (RE), selected regularization parameter at the L-corner (Reg-Para) and time consumption (in second) by the six regularization methods for Fredholm integral equation problem

Method	Tik	TSVD	Land	CGLS	IIL ($p = 3$)	IIE ($p = 3$)
RE	7.62%	7.55%	6.26%	7.00%	6.26%	5.88%
Reg-Para	$\lambda = 0.02535$	$k = 5$	$k = 3242$	$k = 5$	$k = 2583$	$j = 12$
Time	0.27 s	0.20 s	0.27 s	0.16 s	0.20 s	0.19 s

Fig. 5 Filter factors of the six regularization methods under different singular values for Fredholm integral equation problem



methods including the Tikhonov regularization and the TSVD. Also noteworthy is that the regularization effect is guaranteed for the IIE, but may be deviated for the CGLS.

In the above, the IIE is shown to perform well under the prescribed Taylor truncation order $p = 3$. At the closure of this example, the dependence of the performance of the IIE on different Taylor truncation orders p is explored. The detailed results are summarized in Table 3. It is observed that under different truncation orders p , the performance is similar: the relative error is within [5.86%, 6.98%] meaning that all cases give satisfactory solution to the Fredholm integral equation, the time for computation is all near 0.20 s and the number of iterations to arrive at the L-corner is also no more than 13; this is to say, the effect of the order p on the performance of the IIE is not-so-obvious in this example.

4.2 Dynamic force reconstruction of a shear structure

Dynamic force reconstruction of a ten-storey shear structure from the measured acceleration data is considered in this example. The shear structure model is presented in Fig. 6 and the model parameters are taken as: the shear stiffness of every storey $k_1 = k_2 = \dots = k_{10} = 1500$ kN/m and the mass of every floor $m_1 = m_2 = \dots = m_{10} = 3500$ kg. In this way, the stiffness matrix \mathbf{K} and the mass matrix \mathbf{M} are obtained as

$$\mathbf{M} = \begin{pmatrix} m_1 & 0 & \cdots & 0 \\ 0 & m_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & m_{10} \end{pmatrix}, \mathbf{K} = \begin{pmatrix} k_1 + k_2 & -k_2 & \cdots & 0 \\ -k_2 & k_2 + k_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & k_{10} \end{pmatrix}.$$

The Rayleigh damping $\mathbf{C} = a_1 \mathbf{M} + a_2 \mathbf{K}$ is assumed with the first two damping ratios being 1% so that there is $a_1 = 0.0463$ s^{-1} and $a_2 = 1.625 \times 10^{-3}$ s. The structure is initially at rest and a dynamic force with the following expression

$$F(t) = (0.5 + 3e^{-2t} \cos(20t) + \sin(10t)) \times 3500 \text{ N} \quad (63)$$

is enforced at the first floor.

Table 3 Results on relative error (RE), time consumption (in second) and the number of iterations at L-corner by the IIE with different truncation orders p for Fredholm integral equation problem

Truncation order	$p = 1$	$p = 2$	$p = 3$	$p = 5$	$p = 7$
RE	6.98%	6.98%	5.88%	5.86%	5.87%
Time	0.21 s	0.17 s	0.19 s	0.15 s	0.21 s
j at L-corner	$j = 13$	$j = 13$	$j = 12$	$j = 12$	$j = 12$

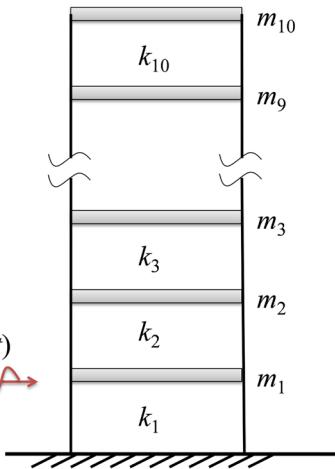


Fig. 6 A ten-storey shear model with dynamic force exerted on the first floor

Herein, the dynamic force is to be inversely reconstructed from the acceleration measurements at the 2nd, 6th and 10th floors. The acceleration data are firstly obtained from numerical simulation at the sampling ratio of 500 Hz within the time duration [0, 2] s and then, the noise of level $e = 2\%$ is added to get the measured data; for a schematic view of the measured data, refer to Fig. 7. Clearly, this force reconstruction problem falls into the general form [40, 41] of Eq. (1): the measured data constitute the right-hand side \mathbf{b} , the force discretized at the same sampling frequency 500 Hz forms the unknown solution \mathbf{x} and the system matrix \mathbf{A} of sizes ($m = 3003, n = 1001$) can be obtained as in the Appendix. Note that it has been suggested from Refs. [40, 41] that for continuous force reconstruction, the matrix \mathbf{L} (see Eq. (3)) for regularization would better be the one in Eq. (4) with $s = 1$; such a choice can enhance the continuity of the solution and is therefore the case in this example. Under this circumstance, the regularization (3) is not in the standard form, nevertheless, the standard form can be easily obtained following the manipulations in Eqs. (5)–(7) with the transformed system matrix $\bar{\mathbf{A}}$ (of size 3003×1000). Note that the conditioning of the force reconstruction problem is determined by the singular values of $\bar{\mathbf{A}}$ and therefore, these singular values are depicted in Fig. 8. It is seen that except for the smallest singular value $\sigma_{1000} = 2.75 \times 10^{-14}$, other singular values lie in the interval [$\sigma_{999} = 5.47 \times 10^{-7}$, $\sigma_1 = 63.34$]. As a result, $\bar{\mathbf{A}}$ is approximately rank-1 deficient and the condition number of $\bar{\mathbf{A}}$ is 2.35×10^{15} , indicating that force reconstruction herein is an ill-posed problem.

The six regularization methods in Table 1 are used to solve this force reconstruction problem where the truncation order takes $p = 3$ for the IIL to differ from the Landweber method, while it is set $p = 1$ for the IIE at first.

Fig. 7 Acceleration data with 2% noise for force reconstruction problem

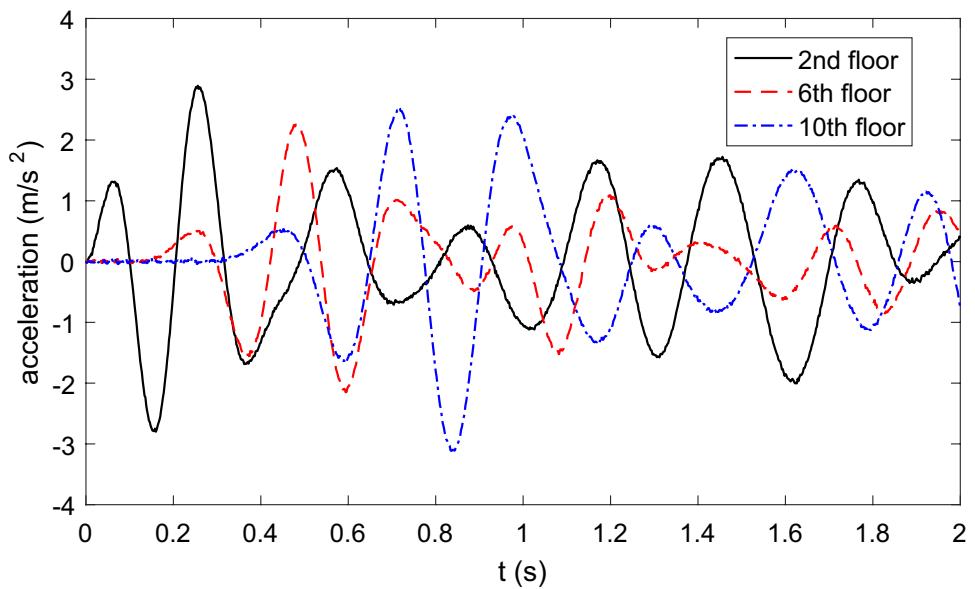
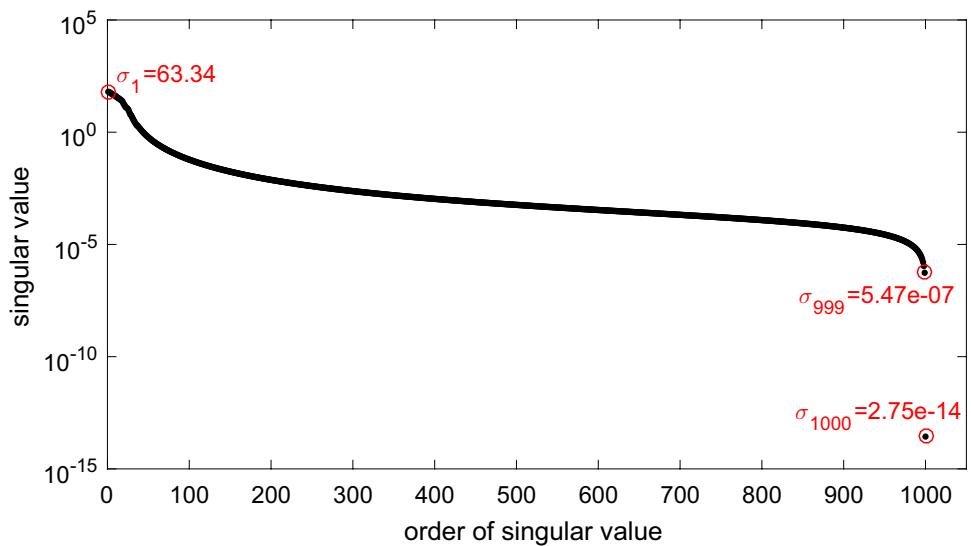


Fig. 8 Schematic of the singular values of the system matrix for force reconstruction problem



The L-curve rule is adopted to determine the regularization parameters and the details are exhibited in Fig. 9. Then, under the selected regularization parameters at the L-corner, the force can be identified and the identification results are displayed in Fig. 10. To further quantify the performance, the relative error (RE) and the computation cost for the six regularization methods are given in Table 4. The above results and further analysis show that

- The L-shape for the log–log scale residual norm vs. solution norm curve is quite apparent (see Fig. 9) especially for the Tikhonov regularization, the TSVD, the CGLS and the IIE. As a result, the selection of the regularization parameter at the L-corner is very straightforward.

- The Landweber method takes 116.33 s to solve the ill-posed problem and the IIL ($p = 3$) takes 39.80 s. These two methods turn out to be much more expensive than other four regularization methods in terms of computation cost. This is mainly because the iterations required to arrive at the L-corner exceed 1000 for the two methods, indicating that very slow convergence is invoked by the Landweber method and the IIL for the ill-posed problem. Thus, the Landweber method and the IIL are often not recommended to solve the ill-posed problem, except that proper preconditioning of the system matrix is introduced to dramatically reduce the condition number; this, however, is beyond the scope of this paper.
- The solution given by the Tikhonov regularization admits the greatest relative error 14.66% so that the

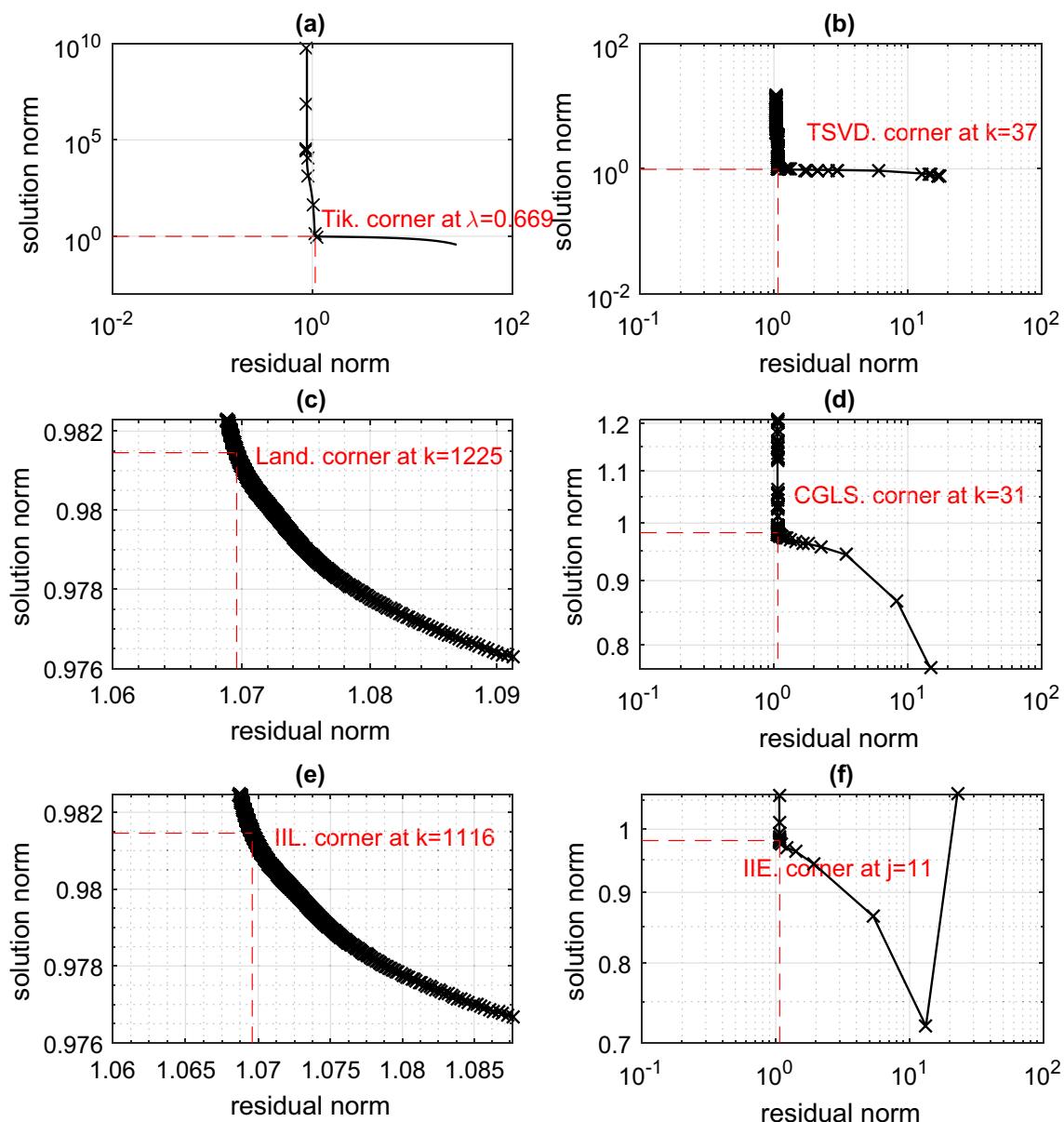


Fig. 9 L-curves of different regularization methods for force reconstruction problem: **a** Tik, **b** TSVD, **c** Land, **d** CGLS, **e** IIL ($p = 3$) and **f** IIE ($p = 1$)

reconstructed force by the Tikhonov regularization deviates most from the true force (see Fig. 10). This may be caused by the not-so-proper selection of the candidate regularization parameters. In this way, the Tikhonov regularization seems to be more cumbersome than other regularization methods because efforts should be made to determine a proper set of candidate regularization parameters so as to proceed the L-curve selection rule. While for other regularization methods, the number of iterations or the truncation order of the SVD is naturally the regularization parameter.

- Except for prohibitive computation cost for the Landweber method and the IIL, the computation cost for other four regularization methods is almost at the same level with the time consumption taken between [3.17 s, 8.67 s]. Among them, the cost for the CGLS and the IIE is less than the cost for the Tikhonov regularization and the TSVD; this shall be benefit from the advantage of iterative regularization methods over the direct regularization methods. It is also seen that the proposed IIE takes less time than the CGLS and the reason may lie in that the

convergence by the IIE is faster (in only $j = 11$ iterations) than by the CGLS (in $k = 31$ iterations).

- Except for the Tikhonov regularization, other five regularization methods have led to satisfactory solution of the force reconstruction problem because the relative errors are all less than 9%, among which the proposed IIE performs best.
- To see the filter effect of the six regularization methods, the filter factors under different singular values are

presented in Fig. 11. All well verify the conditions **P1–P4** in Sect. 2.2, indicating that these six methods can indeed pose the regularization effect for this example.

- To further see the effect of the Taylor truncation order p on the proposed IIE, different ps are selected for the IIE and then, the results are presented in Table 5. It is observed that for different p 's, the number of iterations at the L-corner are the same and the computation cost as well as the relative error is quite near to each other.

Fig. 10 Identified dynamic forces by the six regularization methods for force reconstruction problem

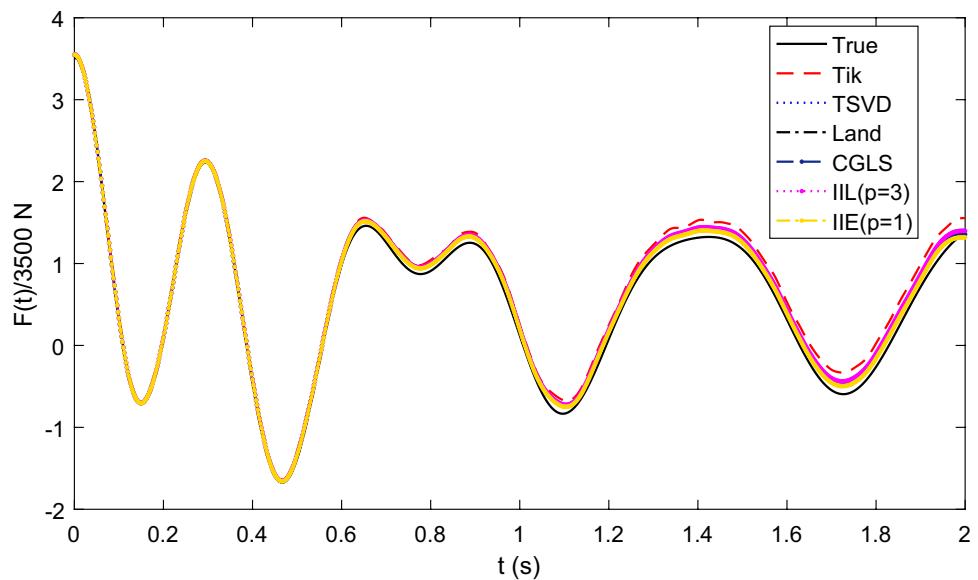


Table 4 Results on relative error (RE), selected regularization parameter at the L-corner (Reg-Para) and time consumption (in second) by the six regularization methods for force reconstruction problem

Method	Tik	TSVD	Land	CGLS	IIL ($p = 3$)	IIE ($p = 1$)
RE	14.66%	7.98%	8.91%	7.06%	8.92%	5.98%
Reg-Para	$\lambda = 0.669$	$k = 37$	$k = 1225$	$k = 31$	$k = 1116$	$j = 11$
Time	6.82 s	8.67 s	116.33 s	6.14 s	39.80 s	3.17 s

Fig. 11 Filter factors of the six regularization methods under different singular values for force reconstruction problem

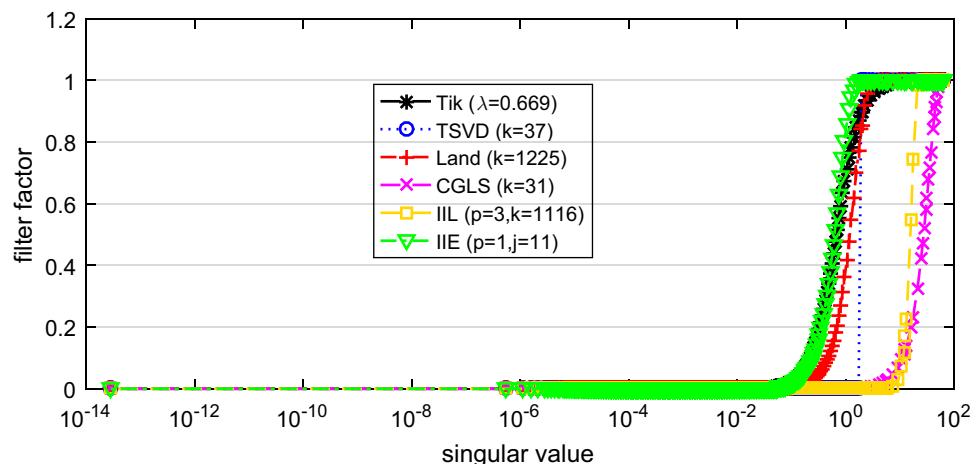


Table 5 Results on relative error (RE), time consumption (in second) and the number of iterations at L-corner by the IIE with different truncation orders p for force reconstruction problem

Truncation order	$p = 1$	$p = 2$	$p = 3$	$p = 5$	$p = 7$
RE	5.98%	5.98%	5.01%	6.08%	5.77%
Time	3.17 s	2.92 s	3.35 s	3.28 s	3.49 s
j at L-corner	$j = 11$				

The performance of the proposed iterative integration method including the IIE is almost independent of the choice of the Taylor truncation order p . The reason for this approximate independence on p is given as follows. The ℓ^2 regularization attempts to filter out the effect of relatively small singular values and to preserve the effect of relatively large singular values. Let the smallest one of the relatively large singular values be σ_s and then, to keep the effect of σ_s , there should hold $1 - (g_p(\tau\sigma_s^2))^{2j} \rightarrow 1$ or $(g_p(\tau\sigma_s^2))^{2j} \rightarrow 0$ for some proper j by referring to the filter functions in Table 1. Note that σ_s is often very small with respect to σ_1 , i.e., $\sigma_s/\sigma_1 \ll 1$, and then, by selecting τ in the way of Remark 3, there is $0 < \tau\sigma_s^2 \ll 1$ and $g_p(\tau\sigma_s^2) \rightarrow 1$. Under this circumstance, the higher order terms in $g_p(\tau\sigma_s^2)$ become negligible and there is $g_p(\tau\sigma_s^2) \approx g_1(\tau\sigma_s^2)$, $p \leq 2$; this just explains why different choices of p often lead to similar regularization performance.

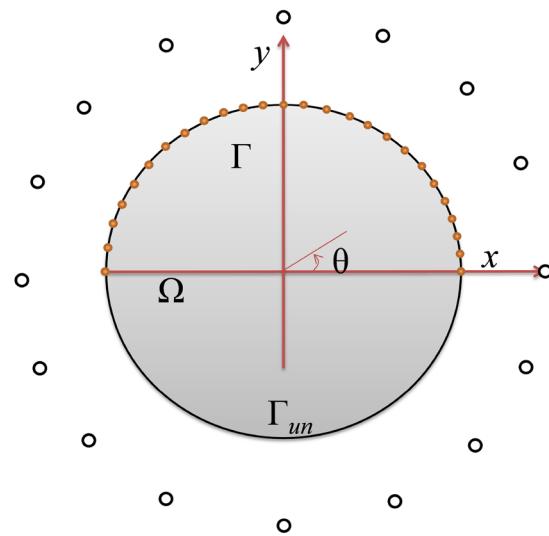
- Following the above point, $1 - g_p(\tau\sigma_s^2) \rightarrow 0$ and then, $(g_p(\tau\sigma_s^2))^k \approx 1 - k(1 - g_p(\tau\sigma_s^2))$ while $(g_p(\tau\sigma_s^2))^{2j} \approx 1 - 2^j(1 - g_p(\tau\sigma_s^2))$ for some proper k, j . This to some extent implies that the IIL admits linear (with respect to k) convergence while the IIE admits exponential (with respect to j) convergence. The proposed IIE is reasonably deduced to converge very fast, e.g., only in 11 iterations for this force reconstruction problem.

4.3 Cauchy problem for steady-state heat conduction

Consider the steady-state heat conduction in a 2D circular domain Ω (see Fig. 12). The conductivity coefficient for this domain is set to unity for simplicity and there is no heat source inside the domain. In this way, the governing equation for steady-state heat conduction is found to be

$$-\Delta u = 0, \text{ in } \Omega \quad (64)$$

where $\mathbf{y} = [x, y] \in \Omega$ and $u = u(\mathbf{y})$ denotes the temperature field. Let Γ be a portion of the whole boundary, i.e., $\Gamma \subset \partial\Omega$



Dots (●): Collection points for Dirichlet and Neumann data
Circles (○): Source points

Fig. 12 Cauchy problem for steady-state heat conduction on a circular domain Ω

and then, the Cauchy problem is described as: given the Dirichlet/temperature and Neumann/heat-flux data on the known boundary Γ with \mathbf{n} being the outward normal vector

$$u|_{\Gamma} = \varphi, \frac{\partial u}{\partial \mathbf{n}}|_{\Gamma} = \psi, \quad (65)$$

determine the boundary condition (or temperature) on the unknown boundary $\Gamma_{un} = \partial\Omega \setminus \Gamma$.

A usual way to solve the Cauchy problem is to use the method of fundamental solutions (MFS) [8]. The basic idea is that, when there is no heat source inside the domain, the temperature inside the domain can be approximately generated by a series of n point heat sources properly locating at $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n$ outside the domain. For a single point heat source at \mathbf{z}_j with unit strength, the resulting temperature is just the fundamental solution (or Green function)

$$G(\mathbf{y}, \mathbf{z}_j) = \frac{1}{2\pi} \ln |\|\mathbf{y} - \mathbf{z}_j\||. \quad (66)$$

Assume that the strengths of the n point heat sources are f_1, f_2, \dots, f_n or $\mathbf{x} = [f_1; f_2; \dots; f_n]$ and this can lead to the following temperature solution

$$u(\mathbf{y}) = \sum_{j=1}^n G(\mathbf{y}, \mathbf{z}_j) f_j. \quad (67)$$

To get the eventual temperature, the strengths \mathbf{x} of the heat sources are to be determined from the temperature and flux data on the known boundary Γ . To go further,

the known boundary data (65) are collected on m points $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m \in \Gamma$, yielding for $i = 1, 2, \dots, m$

$$\begin{aligned} u(\mathbf{y}_i) &= \sum_{j=1}^n G(\mathbf{y}_i, \mathbf{z}_j) f_j = \varphi(\mathbf{y}_i), \\ \frac{\partial u(\mathbf{y}_i)}{\partial \mathbf{n}} &= \sum_{j=1}^n \frac{\partial G(\mathbf{y}_i, \mathbf{z}_j)}{\partial \mathbf{n}} f_j = \psi(\mathbf{y}_i) \end{aligned} \quad (68)$$

and this completely coincides with the linear form in Eq. (1):

$$\mathbf{A}\mathbf{x} = \mathbf{b},$$

$$\begin{aligned} \mathbf{A} &= \begin{pmatrix} \mathbf{A}^{(1)} \\ \mathbf{A}^{(2)} \end{pmatrix}, \mathbf{A}^{(1)} = (G(\mathbf{y}_i, \mathbf{z}_j))_{m \times n}, \mathbf{A}^{(2)} = \left(\frac{\partial G(\mathbf{y}_i, \mathbf{z}_j)}{\partial \mathbf{n}} \right)_{m \times n}, \\ \mathbf{b} &= \begin{pmatrix} \mathbf{b}^{(1)} \\ \mathbf{b}^{(2)} \end{pmatrix}, \mathbf{b}^{(1)} = (\varphi(\mathbf{y}_i))_m, \mathbf{b}^{(2)} = (\psi(\mathbf{y}_i))_m, \end{aligned} \quad (69)$$

where the system matrix \mathbf{A} is of size $2m \times n$.

Herein, the Cauchy problem is specified as follows (refer to Fig. 12): the circular domain Ω is of radius $r = 1$, the known boundary Γ is set as the part of the boundary corresponding to the angle $\theta \in [0, \pi]$, where $m = 101$ points are uniformly selected on this boundary to collect the temperature and flux data, while $n = 100$ source points are uniformly chosen on a circle outside the domain with radius $R = 2$; by these settings, the system matrix \mathbf{A} can be obtained. In this example, the true temperature field is given as

$$u(x, y) = x^3 - 3xy^2 + e^{2y} \sin(2x) - e^x \cos(y) \quad (70)$$

so that the right-hand side \mathbf{b} is obtainable with the noise level $e = 1\%$ (see Eq. (57)). To get more sights into this Cauchy problem, the singular values of the matrix \mathbf{A} are plotted in Fig. 13. It is seen that many singular values approaching 1×10^{-15} cluster at zero and the condition number being 3.04×10^{16} is very huge; this means that the Cauchy problem herein is indeed an ill-posed problem. The four regularization methods including the Tikhonov regularization, the TSVD, the CGLS and the proposed IIE are used to solve this ill-posed Cauchy problem, while other two methods—the Landweber method and the IIL, are excluded because they suffer from very slow convergence. For the IIE, the Taylor truncation order is simply fixed to $p = 1$ due to the fact that the performance of the IIE is shown nearly independent on p . Again, the L-curve rule is applied to determine the regularization parameters where the details are depicted in Fig. 14 and then, after selecting the regularization parameters at the L-corners, the regularized solutions can be obtained.

Note that with the regularized solution obtained, the temperature field over the whole domain Ω is acquired by resorting to Eq. (67), and thereafter, the temperature on the unknown boundary Γ_{un} corresponding to the angle $\theta \in [\pi, 2\pi]$ is reconstructed; to get a schematic impression

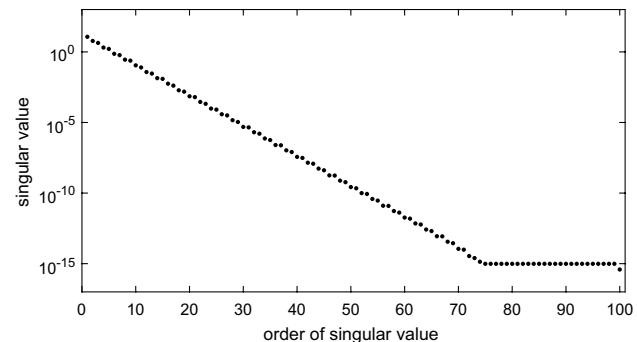


Fig. 13 Schematic of the singular values of the system matrix for Cauchy problem

on the identified temperature on Γ_{un} , refer to Figure 15. As it is observed, all four regularization methods have led to satisfactory reconstruction of the temperature on Γ_{un} and thereof, good solution of the Cauchy problem. To quantify the performance of the four regularization methods, the relative error (RE) between the identified and true temperatures at the 101 uniformly distributed points on Γ_{un} and the time consumption are calculated and listed in Table 6. Though all being satisfactory, the proposed IIE is shown to give most accurate reconstruction of the temperature on Γ_{un} because the RE of the IIE equaling to 3.98% is smallest. As regards the computation cost, the TSVD is found to take least computation time while the Tikhonov regularization takes the most. This is mainly caused by the fact that for this problem, the number of the candidate regularization parameters for the Tikhonov regularization is much greater than the required order $k = 15$ for the TSVD. As with the two iterative regularization methods—the CGLS and the IIE, the consumption time is quite near to each other and the convergence is very fast within $k = 23$ iterations for the CGLS and $j = 19$ iterations for the IIE. Consequently, the effectiveness and efficiency of the proposed IIE for solution of the ill-posed Cauchy problem are verified.

5 Conclusions

A novel iterative integration regularization method has been proposed for robust and efficient solution of ill-posed problems in this paper. It is mainly motivated by the fact that inversion of a positive definite matrix can be alternatively represented in an integral form. Two practical variants—the IIL and the IIE—to compute the integral using the respective linear and exponential iterations are developed. Detailed analysis reveals that the proposed method

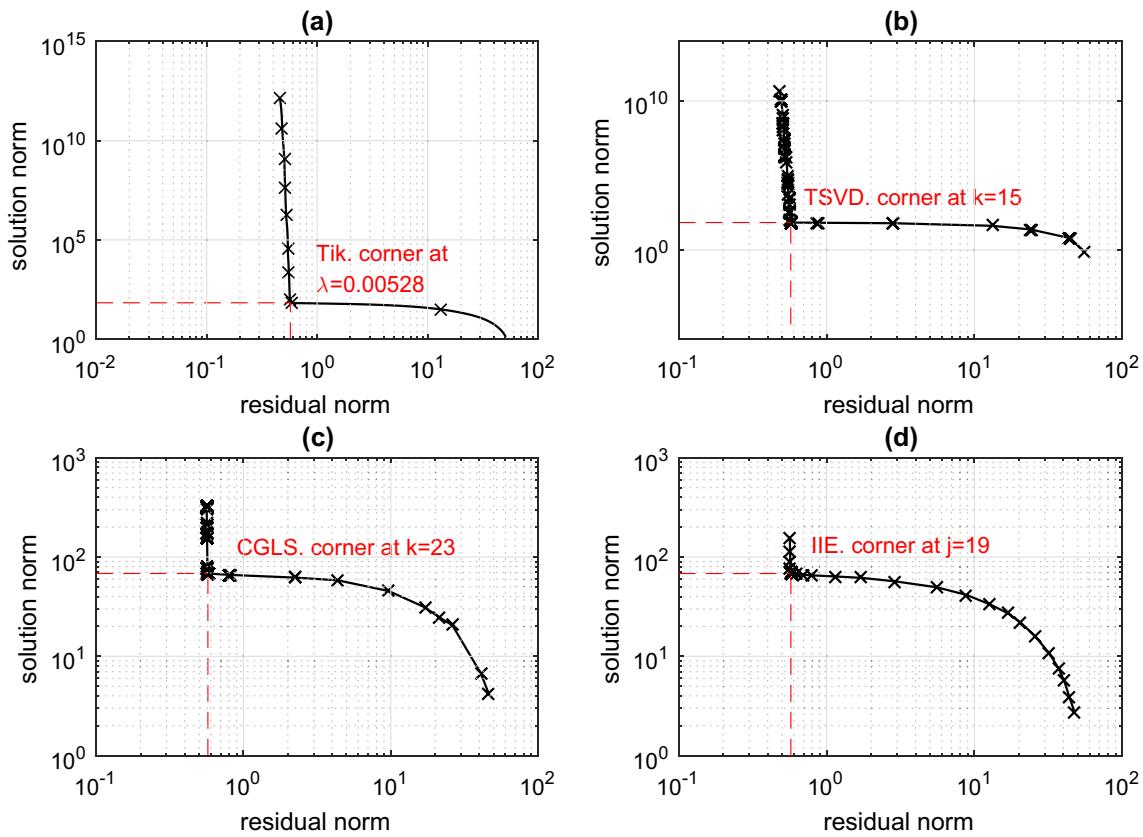


Fig. 14 L-curves of different regularization methods for Cauchy problem: **a** Tik, **b** TSVD, **c** CGLS and **d** IIE ($p = 1$)

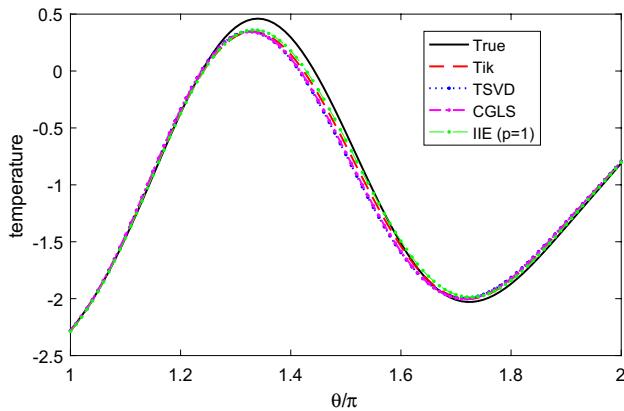


Fig. 15 Identified temperature on the unknown boundary Γ_{un} by different regularization methods for Cauchy problem

possesses the guaranteed regularization effect, i.e., it can filter out the undesired effect of the relatively small singular values, while keep the effect of the relatively large singular values. To verify the performance of the proposed method, three typical ill-posed problems including the Fredholm integral equation, the force reconstruction and the Cauchy problem are studied along with comparison

Table 6 Results on relative error (RE), selected regularization parameter at the L-corner (Reg-Para) and time consumption (in second) by different regularization methods for Cauchy problem

Method	Tik	TSVD	CGLS	IIE ($p = 1$)
RE	5.09%	7.27%	6.77%	3.98%
Reg-Para	$\lambda = 0.00528$	$k = 15$	$k = 23$	$j = 19$
Time	0.296 s	0.078 s	0.140 s	0.109 s

to some usual direct and iterative regularization methods, and results show that

- For the proposed iterative integration regularization method, the exponential iteration (the IIE) is shown in Table 4 to admit much faster convergence and more efficient computation than the linear iteration (the IIL) and therefore, the IIE is preferred in the proposed iterative integration method.
- The proposed IIE is found able to effectively and efficiently solve all the three ill-posed problems.
- Compared to general direct regularization methods including the Tikhonov regularization and the TSVD, the pro-

posed IIE usually turns out to be more efficient by referring to the results in Table 4; this is mainly due to the iterative nature of the proposed IIE.

- The Landweber method, as an iterative regularization method and a particular case of the IIL, is shown to admit very slow convergence and thereof, prohibitively high computation cost for ill-posed problems. In other words, the proposed IIE is much more efficient than the Landweber method.
- Focusing on the proposed IIE and the CGLS, both methods are iterative; however, the IIE is stationary while the CGLS is non-stationary whose filter factors depend on the right-hand side. In all the three examples, the relative error and the computation cost for the CGLS and the IIE are at the same level, that is to say, the IIE can perform as well as the CGLS. Nevertheless, the main advantage of the proposed IIE over the CGLS resides at that the IIE has the guaranteed regularization effect, while the CGLS may not [25].

Acknowledgements The present investigation was performed under the support of National Natural Science Foundation of China (nos. 11702336 and 11972380), Guangdong Province Natural Science Foundation (no. 2017A030313007 and no. 2018B030311001) and the Fundamental Research Funds of the Central Universities (no. 17lgpy54 and no. 19lgpy106).

$$\mathbf{b} = \mathbf{Ax}$$

$$\mathbf{b} = [\mathbf{d}(t_1); \mathbf{d}(t_2); \dots; \mathbf{d}(t_{N_T})], \mathbf{x} = [\mathbf{f}(t_1); \mathbf{f}(t_2); \dots; \mathbf{f}(t_{N_T})]$$

$$\mathbf{A} = \begin{pmatrix} \mathbf{DG} & & \\ \mathbf{OZ}_1\mathbf{G} & \mathbf{DG} + \mathbf{OZ}_2\mathbf{G} & \\ \vdots & \vdots & \ddots \\ \mathbf{OB}^{N_T-3}\mathbf{Z}_1\mathbf{G} & \mathbf{OB}^{N_T-4}\mathbf{Z}_1\mathbf{G} + \mathbf{OB}^{N_T-3}\mathbf{Z}_2\mathbf{G} & \dots & \mathbf{DG} + \mathbf{OZ}_2\mathbf{G} \\ \mathbf{OB}^{N_T-2}\mathbf{Z}_1\mathbf{G} & \mathbf{OB}^{N_T-3}\mathbf{Z}_1\mathbf{G} + \mathbf{OB}^{N_T-2}\mathbf{Z}_2\mathbf{G} & \dots & \mathbf{OZ}_1\mathbf{G} + \mathbf{OBZ}_2\mathbf{G} & \mathbf{DG} + \mathbf{OZ}_2\mathbf{G} \end{pmatrix}. \quad (75)$$

Appendix

To get the matrix \mathbf{A} for force reconstruction in structural systems, the Newmark method for solution of the dynamic equation is called for $k = 1, 2, \dots, N_T$,

$$\begin{cases} \mathbf{M}\ddot{\mathbf{u}}(t_{k+1}) + \mathbf{C}\dot{\mathbf{u}}(t_{k+1}) + \mathbf{K}\mathbf{u}(t_{k+1}) = \mathbf{G}\mathbf{f}(t_{k+1}), \\ \dot{\mathbf{u}}(t_{k+1}) = \dot{\mathbf{u}}(t_k) + [\gamma\ddot{\mathbf{u}}(t_{k+1}) + (1 - \gamma)\dot{\mathbf{u}}(t_k)]\Delta t, \\ \mathbf{u}(t_{k+1}) = \mathbf{u}(t_k) + \dot{\mathbf{u}}(t_k)\Delta t + [\beta\ddot{\mathbf{u}}(t_{k+1}) + (1/2 - \beta)\dot{\mathbf{u}}(t_k)]\Delta t^2, \end{cases} \quad (71)$$

where $\gamma = 1/2, \beta = 1/4$ are fixed in this paper, $\mathbf{M}, \mathbf{C}, \mathbf{K}$ are the mass, damping and stiffness matrices, $\mathbf{u}(t)$ denotes the displacements, $\mathbf{f}(t)$ collects all independent forces and \mathbf{G} describes the spatial distributions of the forces. Simple

manipulations on Eq. (71) can yield the following state equation

$$\begin{pmatrix} \mathbf{u}(t_{k+1}) \\ \dot{\mathbf{u}}(t_{k+1}) \end{pmatrix} = \mathbf{B} \begin{pmatrix} \mathbf{u}(t_k) \\ \dot{\mathbf{u}}(t_k) \end{pmatrix} + \mathbf{Z}_1\mathbf{G}\mathbf{f}(t_k) + \mathbf{Z}_2\mathbf{G}\mathbf{f}(t_{k+1}), \quad (72)$$

with

$$\begin{aligned} \mathbf{B} &= \begin{pmatrix} \gamma\mathbf{K} & \gamma\mathbf{C} + \mathbf{M}/\Delta t \\ 2\beta\mathbf{K} + 2\mathbf{M}/\Delta t^2 & 2\beta\mathbf{C} \end{pmatrix}^{-1} \\ &\quad \begin{pmatrix} -(1 - \gamma)\mathbf{K} & -(1 - \gamma)\mathbf{C} + \mathbf{M}/\Delta t \\ -(1 - 2\beta)\mathbf{K} + 2\mathbf{M}/\Delta t^2 & -(1 - 2\beta)\mathbf{C} + 2\mathbf{M}/\Delta t \end{pmatrix} \\ \mathbf{Z}_1 &= \begin{pmatrix} \gamma\mathbf{K} & \gamma\mathbf{C} + \mathbf{M}/\Delta t \\ 2\beta\mathbf{K} + 2\mathbf{M}/\Delta t^2 & 2\beta\mathbf{C} \end{pmatrix}^{-1} \begin{pmatrix} (1 - \gamma)\mathbf{I} \\ (1 - 2\beta)\mathbf{I} \end{pmatrix} \\ \mathbf{Z}_2 &= \begin{pmatrix} \gamma\mathbf{K} & \gamma\mathbf{C} + \mathbf{M}/\Delta t \\ 2\beta\mathbf{K} + 2\mathbf{M}/\Delta t^2 & 2\beta\mathbf{C} \end{pmatrix}^{-1} \begin{pmatrix} \gamma\mathbf{I} \\ 2\beta\mathbf{I} \end{pmatrix} \end{aligned} \quad (73)$$

where \mathbf{I} denotes the identity matrix, $\Delta t = t_{k+1} - t_k$. The measured data $\mathbf{d}(t_k), \forall k$ are often of the form

$$\mathbf{d}(t_k) = \mathbf{O} \begin{pmatrix} \mathbf{u}(t_k) \\ \dot{\mathbf{u}}(t_k) \end{pmatrix} + \mathbf{D}\mathbf{G}\mathbf{f}(t_k), \quad (74)$$

where \mathbf{O}, \mathbf{D} are observation matrices. Setting the initial states $\mathbf{u}(t_1) = \dot{\mathbf{u}}(t_1) = \mathbf{0}$ and combining (72) with (74), it is acquired that

References

1. Ertveldt J, Pintelon R, Vanlanduit S (2016) Identification of unsteady aerodynamic forces from forced motion wind tunnel experiments. AIAA J 54(10):3265–3273
2. Maes K, Weijtjens W, de Ridder EJ, Lombaert G (2018) Inverse estimation of breaking wave loads on monopile wind turbines. Ocean Eng 163:544–554
3. Wang L, Liu J, Xie Y, Gu Y (2018) A new regularization method for the dynamic load identification of stochastic structures. Comput Math Appl 76:741–759
4. Xu H, Zhang L, Li Q (2019) A novel inverse procedure for load identification based on improved artificial tree algorithm. Eng Comput. <https://doi.org/10.1007/s00366-019-00848-4>
5. Ahmadian H, Mottershead JE, Friswell MI (1998) Regularization methods for finite element model updating. Mech Syst Signal Process 12(1):47–64
6. Titurus B, Friswell MI (2008) Regularization in model updating. Int J Numer Methods Eng 75:440–478

7. Rubio PB, Louf F, Chamoin L (2018) Fast model updating coupling Bayesian inference and PGD model reduction. *Comput Mech* 62(6):1485–1509
8. Wei T, Hon YC, Ling L (2007) Method of fundamental solutions with regularization techniques for Cauchy problems of elliptic operators. *Eng Anal Bound Elem* 31:373–385
9. Shivanian E, Jafarabadi A (2017) Inverse Cauchy problem of annulus domains in the framework of spectral meshless radial point interpolation. *Eng Comput* 33:431–442
10. Kabannikh SI (2008) Definitions and examples of inverse and ill-posed problems. *J Inverse Ill Posed Probl* 16:317–357
11. Benning M, Burger M (2018) Modern regularization methods for inverse problems. *Acta Numer* 27:1–111
12. Tikhonov AN (1963) On the solution of ill-posed problems and the method of regularization. *Soviet Mathe* 4:1035–1038
13. Hansen PC (1992) Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Rev* 34(4):561–580
14. Bryan K, Leise T (2013) Making do with less: an introduction to compressed sensing. *SIAM Rev* 55(3):547–566
15. Kim SJ, Koh K, Boyd S, Gorinevsky D (2009) ℓ_1 trend filtering. *SIAM Rev* 51(2):339–360
16. Grasmair M, Haltmeier M, Scherzer O (2008) Sparse regularization with ℓ^q penalty term. *Inverse Probl* 24(5):055020
17. Strong D, Chan T (2003) Edge-preserving and scale-dependent properties of total variation regularization. *Inverse Probl* 19:S165–S187
18. Hansen PC, O’Leary DP (1993) The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM J Sci Comput* 14(6):1487–1503
19. Hansen PC (2007) Regularization tools version 4.1 for matlab 7.3. *Numer Algorithms* 46:189–194
20. Hansen PC (1987) The truncated SVD as a method for regularization. *BIT Numer Math* 27:354–553
21. Hansen PC (2002) Deconvolution and regularization with Toeplitz matrices. *Numer Algorithms* 29:323–378
22. Wang L, Xie Y, Wu Z, Du Y, He K (2019) A new fast convergent iteration regularization method. *Eng Comput* 35:127–138
23. Landweber L (1951) An iteration formula for Fredholm integration equations of the first kind. *Am J Math* 73:615–624
24. Hanke M (1995) The minimal error conjugate gradient method is a regularization method. *Proc Am Math Soc* 123(11):3487–3497
25. Nagy JG, Palmer KM (2003) Steepest descent, CG, and iterative regularization of ill-posed problems. *BIT Numer Math* 43:1003–1017
26. Bauer F, Lukas MA (2011) Comparing parameter choice methods for regularization of ill-posed problems. *Math Comput Simul* 81:1795–1841
27. Lukas MA (2006) Robust generalized cross-validation for choosing the regularization parameter. *Inverse Probl* 22:1883–1902
28. Liao H, Ng MK (2011) Blind deconvolution using generalized cross-validation approach to regularization parameter estimation. *IEEE Trans Image Process* 20(3):670–680
29. Hou R, Xia Y, Bao Y, Zhou X (2018) Selection of regularization parameter for ℓ_1 -regularized damage detection. *J Sound Vib* 423:141–160
30. Rieder A (2005) Inexact Newton regularization using conjugate gradients as inner iteration. *SIAM J Numer Anal* 43(2):604–622
31. Hanke M (2001) On Lanczos based methods for the regularization of discrete ill-posed problems. *BIT Numer Math* 41(5):1008–1018
32. Marin L, Lesnic D (2005) The method of fundamental solutions for the Cauchy problem associated with two-dimensional Helmholtz-type equations. *Comput Struct* 83:267–278
33. Dongarra J, Gates M, Haidar A, Kurzak J, Luszczek P, Tomov S, Yamazaki I (2018) The singular value decomposition: anatomy of optimizing an algorithm for extreme scale. *SIAM Rev* 60(4):808–865
34. Berntsson F, Lin C, Xu T, Wokiyi D (2017) An efficient regularization method for a large scale ill-posed geothermal problem. *Comput Geosci* 105:1–9
35. Hansen PC (1998) Rank-deficient and discrete ill-posed problems. Numerical aspects of linear inversion. SIAM, Philadelphia
36. Zhong WX (2004) On precise integration method. *J Comput Appl Math* 163:59–78
37. Castellanos JL, Gomez S, Guerra V (2002) The triangle method for finding the corner of the L-curve. *Appl Numer Math* 43:359–373
38. Wazwaz AM (2011) The regularization method for Fredholm integral equations of the first kind. *Comput Math Appl* 61:2981–2986
39. Vogel CR, Oman ME (1998) Fast, robust total variation-based reconstruction of noisy, blurred images. *IEEE Trans Image Process* 7(6):813–824
40. Chen Z, Chan THT (2017) A truncated generalized singular value decomposition algorithm for moving force identification with ill-posed problems. *J Sound Vib* 401:297–310
41. Chen Z, Chan THT, Nguten A (2018) Moving force identification based on modified conjugate gradient method. *J Sound Vib* 423:100–117

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.