Solving Severely Ill-Posed Linear Systems with Time Discretization Based Iterative Regularization Methods

GONG Rongfang, *HUANG Qin**

College of Science, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, P. R. China

Abstract: Recently, inverse problems have attracted more and more attention in computational mathematics and become increasingly important in engineering applications. After the discretization, many of inverse problems are reduced to linear systems. Due to the typical ill-posedness of inverse problems, the reduced linear systems are often ill-posed, especially when their scales are large. This brings great computational difficulty. Particularly, a small perturbation in the right side of an ill-posed linear system may cause a dramatical change in the solution. Therefore, regularization methods should be adopted for stable solutions. In this paper, a new class of accelerated iterative regularization methods is applied to solve this kind of large-scale ill-posed linear systems. An iterative scheme becomes a regularization method only when the iteration is early terminated. And a Morozov's discrepancy principle is applied for the stop criterion. Compared with the conventional Landweber iteration, the new methods have acceleration effect, and can be compared to the well-known accelerated ν-method and Nesterov method. From the numerical results, it is observed that using appropriate discretization schemes, the proposed methods even have better behavior when comparing with ν-method and Nesterov method.

Key words: linear system; ill-posedness; large-scale; iterative regularization methods; acceleration

CLC number: O241   **Document code:** A   **Article ID:** 1005-1120(2020)06-0979-16

0 Introduction

During the past fifty years, inverse problems have attracted more and more attention and have extensive applications in engineering and mathematical fields, such as Cauchy problem[1-3], geophysical exploration[4], steady heat conduction problems[5-6], inverse scattering problems[7-9], image processing[10-11], and reconstructon of radiated noise source[12] etc. After the discretization, many of them are reduced to the ill-posed linear system as

$$Ax = b \tag{1}$$

Here by the ill-posed linear system, it means that the condition number of the coefficient matrix $A \in \mathbf{R}^{m \times n}$ is much larger compared with the square of the scale of A. Denote by the quasi-solution of Eq.(1). There are three cases for $x^{\dagger}$: (i) The system admits a unique solution, then $x^{\dagger}$ is the exact solution; (ii) the system has more than one solution, then $x^{\dagger}$ is the one of minimal 2-norm among all solutions; (iii) the system has no solution, then $x^{\dagger}$ is least squares solution. It is easy to verify that $x^{\dagger}$ is unique.

Practically, b often comes from measurements or discretization, and contains inevitably noise. Instead of b, assume we only have noisy data b^{δ} at hand, satisfying

$$\| b - b^{\delta} \| \leqslant \delta \tag{2}$$

where δ is the noisy level and $\|\cdot\|$ the standard Euclidean norm of a vector. Therefore, in this paper we are devoted to find an approximate solution to the polluted linear system

$$Ax = b^{\delta} \tag{3}$$

Similarly, denote by x^{δ} the unique quasi-solu-

tion corresponding to the noisy data b^δ.

For an ill-posed system, a small disturbance in b will lead to much large change in solution x. This brings quite large difficulty when one solves the system numerically, especially when the scale is large. Thus, it is useless to use the conventional numerical methods to solve system (1) or (3). In fact, by using the singular value decomposition (SVD), the solutions of Eq.(1) or (3) can be written formally as[13]

$$x^\dagger = \sum_{i=1}^{r} \frac{u_i^T b}{\sigma_i} v_i, x^\delta = \sum_{i=1}^{r} \frac{u_i^T b^\delta}{\sigma_i} v_i$$

where $\{(\sigma_i, u_i, v_i)\}$ are the singular values and singular vectors of the coefficient matrix A, satisfying $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r > 0$, $\sigma_i \leqslant 1$ for $i \geqslant k$ and some index $k \leqslant r$, r is the rank of A. Therefore $\|x^\delta - x^\dagger\|$ may be quite big even if $\|b^\delta - b\|$ is small, and regularization methods are needed for obtaining a stable approximate solution to x^δ and thus also to $x^\dagger$.

Generally speaking, there are three groups of regularization methods: truncated singular value decomposition (TSVD)[13-15], variational regularization methods[16-17] and iterative regularization methods[18-21]. For large-scale ill-posed problems, TSVD leads to computational difficulty in that the vast storage and the heavy computing burden. The most famous variational regularization method is Tikhonov's regularization method[22]. Recently, a projection fractional Tikhonov regularization method is proposed[23]. However, during the determination of the regularization parameters, a forward problem has to be solved for each regularization parameter, which makes the calculation very large. Iterative regularization methods have the advantages of low computational cost and simple forms. Thus for large-scale problem, we prefer to use iterative regularization methods, where iterative schemes are proposed to solve the following optimization problem.

$$\min J(x) = \frac{1}{2} \|Ax - b^\delta\|^2 \qquad (4)$$

The most classic iterative regularization method should be Landweber iteration[24-25]. For the linear system (3), Landweber iteration is defined by

$$x_{k+1} = x_k - \Delta t \nabla J(x_k) \qquad (5)$$

where $\nabla J(x_k) = A^T(Ax_k - b^\delta)$ and $\Delta t \in (0, 2/\|A\|^2)$.

Eq.(5) can be viewed as a discrete analog of

the following first order evolution equation.

$$\begin{cases} \dot{x}(t) = -\nabla J(x(t)) \\ x(0) = x_0 \end{cases} \qquad (6)$$

where t is the introduced artificial time, and x_0 an initial guess to $x^\dagger$. The formulation (6) is known as an asymptotical regularization, or the Showalter's method[26-27]. It is well known that the Landweber method converges quite slowly[28]. Thus, it is no wonder that accelerating strategies are adopted in practice. In recent years, there has been increasing evidence to show that the second order iterative methods exhibit remarkable acceleration properties for ill-posed problems. The most well-known methods are the Levenberg-Marquardt method[29], the iteratively regularized Gauss-Newton method[30], the Nesterov acceleration scheme[31] and the ν-method[28]. Here, the iterative schemes are proposed by discretizing the following second order evolution system[32].

$$\begin{cases} \ddot{x}(t) + \eta(t)\dot{x}(t) + A^T Ax(t) = A^T b^\delta \quad t \in (t_0, \infty) \\ x(t_0) = x_0, \dot{x}(t_0) = \dot{x}_0 \end{cases}$$
$$(7)$$

where $(x_0, \dot{x}_0)$ is the prescribed initial data, $\eta > 0$, uniformly bounded, is the damping parameter, and may or may not depends on the time t. Note that the system (7) becomes a regularization method only when the evolution process is stopped in advance. Denote by T the terminating time. Then it plays the role of the regularization parameter and should be chosen properly. Here, the Morozov's discrepancy principle is as the stopping criterion

$$\|Ax(T) - b^\delta\| \leqslant \tau\delta \qquad (8)$$

where τ is a fixed positive number. Like Landweber method, the well-known Nesterov acceleration method can be viewed as the discrete analog of Eq.(7). In fact, for all fixed $T > 0$[27]

$$\lim_{\omega \to 0} \max_{0 \leqslant k \leqslant T/\sqrt{\omega}} \|x_k - x(k\sqrt{\omega})\| = 0$$

where $x(\cdot)$ is the solution of Eq.(7) with $\eta(t) = \alpha/t$, and $\{x_k\}$ the sequence, generated by the Nesterov's scheme with parameters (α, ω).

The purpose of this paper is to apply the second order dynamical system (7), together with the discrepancy principle (8) for the choice of the termi-

nating time, to the ill-posed linear system (3). Three aspects are expected to be addressed: (i) On one hand, unlike the existing work, where the exact solution is often assumed to exist, the system here may have no solution, and thus the existing theoretical results could not be used directly; on the other hand, benefiting by the linearity and the finite dimension of the problem, compared with the theoretical analysis in the literature, arguments here can be largely simplified. (ii) Effect of the magnitude of the problem on the iterations of different methods is investigated. (iii) Effect of the ill-posedness extent of the problem on the iterations of different methods is discussed.

1　Theoretical Analysis of Continuous Second-Order Flow

In this section, we are devoted to give a series of theoretical analysis on the second order dynamical system (7). Without loss of generality, we set $t_0 = 0$ when $\eta = $ constant, and $t_0 = 1$ when η is time-dependent; set $\dot{x}_0 = 0$ when η is time-dependent. Moreover, for a dynamical damping parameter, we take $\eta = (1 + 2s)/t$ (the constant $s > -1/2$) as an example for the theoretical analysis. Theoretical results for other choice such as $\eta = 1/\sqrt{t}$ could be analyzed similarly, and we omit here.

The definition of a regularized solution is first introduced.

Definition　Let $x(t) \in \mathbf{R}^n$ be the solution of System (7). Then, $x(T^\delta)$, equipped with an appropriate terminating time $T^\delta = T(\delta, b^\delta)$, is called a second order asymptotical regularization solution of Eq.(7) if $x(T^\delta)$ converges (strongly) to $x^\dagger$ in $\mathbf{R}^n$ as $\delta \to 0$.

1.1　Existence and uniqueness of solution trajectory

About the global existence and uniqueness of solution to Eq.(7), the following results can be obtained.

Theorem 1　For any pair $(x_0, \dot{x}_0) \in \mathbf{R}^n \times \mathbf{R}^n$, there exists a unique solution $x(\cdot) \in C^\infty([t_0, \infty), \mathbf{R}^n)$ for the second order dynamical system (7). More-

over, x depends continuously on the data b^δ.

Proof　Denote $z = (x, \dot{x})^\mathrm{T}$, $z_0 = (x_0, \dot{x}_0)^\mathrm{T}$ and rewrite Eq.(7) as a first order differential system.

$$\dot{z}(t) = Bz(t) + \bar{b} \triangleq F(z(t)) \qquad (9)$$

where $B = \begin{pmatrix} 0 & I \\ -A^\mathrm{T}A & -\eta(t)I \end{pmatrix}$, $\bar{b} = \begin{pmatrix} 0 \\ A^\mathrm{T}b^\delta \end{pmatrix}$ and I

denotes the identity matrix of order n. Since A is a bounded matrix and η is uniformly bounded, B is a uniformly bounded matrix. Hence, $F(\cdot)$ is a global Lipschitz functional. By the Cauchy-Lipschitz-Picard theorem[33-34], the first order autonomous system (9) has a global unique solution z for any given initial data z_0, and then the second order dynamical system (7) has a global unique solution x for any given initial data $(x_0, \dot{x}_0)^\mathrm{T}$. Since $A^\mathrm{T}(b^\delta - Ax)$ is linear with respect to x, and $\eta = $ constant or $(1 + 2s)/t$, both of them are infinitely differentiable, which gives $x(\cdot) \in C^\infty([t_0, \infty), \mathbf{R}^n)$. Moreover, the continuous dependence of x on the data b^δ can be easily verified.

Next the relationship between the solution $x(t)$ of Eq.(7) and the exact one $x^\dagger$ is shown. For the future use, denote by S the set of minimizers of $J(\cdot)$ which can be characterized by

$$S = \{ x \in \mathbf{R}^n | \nabla J(x) = 0 \}$$

It is easy to conclude that S is a non-void closed convex subset of $\mathbf{R}^n$. Moreover, for the statement of clarity, the discussion is divided into two parts, corresponding to the noise-free data b and noisy data b^δ respectively.

1.2　Limiting behavior of the solution for noise-free data

Define the modified energy functional of x as

$$E(t) = J(x(t)) - J(x^\dagger) + \frac{1}{2}\|\dot{x}(t)\|^2 \quad (10)$$

Case I　$\eta = $ constant and $t_0 = 0$, that is, we consider the following system

$$\begin{cases} \ddot{x}(t) + \eta\dot{x}(t) + A^\mathrm{T}Ax(t) = A^\mathrm{T}b & t \in (0, \infty) \\ x(0) = x_0, \dot{x}(0) = \dot{x}_0 \end{cases} \quad (11)$$

Proposition 1　Let x be the solution of Eq.(11). Then, the following properties hold

(i) $x(\cdot) \in L^\infty([0, \infty), \mathbf{R}^n)$

(ii) $\dot{x}(\cdot) \in L^\infty([0, \infty), \mathbf{R}^n) \cap L^2([0, \infty), \mathbf{R}^n)$

and thus

$$\dot{x}(t) \to 0 \text{ as } t \to \infty$$

(iii) $\ddot{x}(\cdot) \in L^{\infty}([0, \infty), \mathbf{R}^{n})$ and $\ddot{x}(t) \to 0$ as $t \to \infty$

(iv) $\lim_{t \to \infty} \nabla J(x(t)) = 0$

(v) $J(x(t)) - J(x^{\dagger}) = o(t^{-1})$ as $t \to \infty$

Proof (i) Differentiating $E(t)$ of Eq.(10) and using the system (11) to give

$$\dot{E}(t) = -\eta \|\dot{x}(t)\|^{2} \qquad (12)$$

which means $E(t)$ is non-increasing, and thus

$$E(t) \leqslant E(0) \qquad (13)$$

holds for all $t \geqslant 0$. Consequently, $x(\cdot) \in L^{\infty}([0, \infty), \mathbf{R}^{n})$ by combining Eq.(13) and the coerciveness of $J(\cdot)$, i.e.

$$\lim_{\|x\| \to +\infty} J(x) = +\infty$$

(ii) On one hand, due to Eq.(13), we have

$$\|\dot{x}(t)\|^{2} \leqslant 2(E(t) + J(x^{\dagger})) \leqslant 2(E(0) + J(x^{\dagger}))$$

which gives $\dot{x}(\cdot) \in L^{\infty}([0, \infty), \mathbf{R}^{n})$. On the other hand, by using the decrease and non-negativity of $E(t)$, the limit $E_{\infty} := \lim_{t \to \infty} E(t)$ exists. Therefore, integrating both sides in Eq.(12), we obtain

$$\int_{0}^{\infty} \|\dot{x}(t)\|^{2} \mathrm{d}t = \frac{E(0) - E_{\infty}}{\eta} < \infty$$

which yields $\dot{x}(\cdot) \in L^{2}([0, \infty), \mathbf{R}^{n})$.

In addition, according to a classical result, $\dot{x}(\cdot) \in L^{\infty}([0, \infty), \mathbf{R}^{n}) \cap L^{2}([0, \infty), \mathbf{R}^{n})$ implies $\dot{x}(t) \to 0$ as $t \to \infty$.

(iii) From Eq.(11), we obtain

$$\ddot{x}(t) = -\eta \dot{x}(t) + A^{\mathrm{T}}(b - Ax(t))$$

which gives immediately that $\ddot{x}(\cdot) \in L^{\infty}([0, \infty), \mathbf{R}^{n})$.

By differentiating the first equation of Eq.(11), we obtain

$$\dddot{x}(t) + \eta \ddot{x}(t) = -A^{\mathrm{T}}A\dot{x}(t) \triangleq g(t) \qquad (14)$$

Denote by $y = \ddot{x}$. Then Eq.(14) is reduced to

$$\dot{y}(t) + \eta y(t) = g(t)$$

which implies $y \to 0$ as $t \to \infty$ by noticing that $\eta > 0$ and $g(t) \to 0$ as $t \to \infty$. Thus $\ddot{x}(t) \to 0$ as $t \to \infty$.

(iv) By using Eq.(11) again, we have

$$\nabla J(x(t)) = A^{\mathrm{T}}(Ax(t) - b) = -\ddot{x}(t) - \eta \dot{x}(t)$$

which gives $\lim_{t \to \infty} \nabla J(x(t)) = 0$ by using the facts in properties (ii) and (iii).

(v) Define

$$h(t) = \frac{\eta}{2}\|x(t) - x^{\dagger}\|^{2} + \langle \dot{x}(t), x(t) - x^{\dagger} \rangle \quad (15)$$

By elementary calculations, we derive that

$$\dot{h}(t) = \eta \langle \dot{x}(t), x(t) - x^{\dagger} \rangle + \langle \ddot{x}(t), x(t) - x^{\dagger} \rangle + \|\dot{x}(t)\|^{2} = \|\dot{x}(t)\|^{2} - \|Ax(t) - b\|^{2}$$

which implies that (by noting $E(t) = -\eta\|\dot{x}(t)\|^{2}$)

$$\dot{E}(t) + \eta E(t) + \frac{\eta}{2}\dot{h}(t) = 0$$

or

$$\eta E(t) = -\dot{E}(t) - \frac{\eta}{2}\dot{h}(t)$$

Integrate the above equation on $[0, T]$ to obtain, together with the non-negativity of $E(t)$

$$\int_{0}^{T} E(t)\mathrm{d}t = \frac{1}{\eta}(E(0) - E(T)) + \frac{1}{2}(h(0) - h(T)) \leqslant \frac{1}{\eta}E(0) + \frac{1}{2}(h(0) - h(T)) \quad (16)$$

Due to properties (i)—(ii), $x(t), \dot{x}(t)$ are uniformly bounded, and so is $h(t)$. Hence, letting $T \to \infty$ in Eq.(16), we obtain

$$\int_{0}^{\infty} E(t)\mathrm{d}t < \infty \qquad (17)$$

Since $E(t)$ is non-increasing, we deduce that

$$\int_{T/2}^{T} E(t)\mathrm{d}t \geqslant \frac{T}{2}E(T) \geqslant \frac{T}{2}(J(x(T)) - J(x^{\dagger})) \qquad (18)$$

Using Eq.(17), the left side of Eq.(18) tends to 0 when $T \to \infty$, which implies that $\lim_{t \to \infty} t(J(x(t)) - J(x^{\dagger}))/2 = 0$ or $J(x(t)) - J(x^{\dagger}) = o(t^{-1})$ as $t \to \infty$. The proof is completed.

Case II $\eta(t) = (1 + 2s)/t$ and $t_{0} = 1$

Now, we consider the following evolution system

$$\begin{cases} \ddot{x}(t) + \dfrac{1 + 2s}{t}\dot{x}(t) + A^{T}Ax(t) = A^{T}b & t \in (1, \infty) \\ x(1) = x_{0}, \dot{x}(1) = 0 \end{cases}$$

$$(19)$$

Proposition 2 Let x be the solution of Eq.(19). Then, for $s \geqslant 1$, the following properties hold

(i) $x(\cdot) \in L^{\infty}([1, \infty), \mathbf{R}^{n})$

(ii) $\dot{x}(\cdot) \in L^{\infty}([1, \infty), \mathbf{R}^{n}) \cap L^{2}([1, \infty), \mathbf{R}^{n})$ and thus $\dot{x}(t) \to 0$ as $t \to \infty$

(iii) $\ddot{x}(\cdot) \in L^{\infty}([1, \infty), \mathbf{R}^{n})$

(iv) $J(x(t)) - J(x^{\dagger}) = \mathcal{O}(t^{-2})$ as $t \to \infty$

Proof (i) Differentiating $E(t)$ of Eq.(10)

and using the system (19) to give

$$\dot{E}(t) = -\frac{1+2s}{t}\|\dot{x}(t)\|^2 \qquad (20)$$

The remaining proof for property (i) is similar to that for property (i) of proposition 1.

(ii) Like the statements in the proof of proposition 1, we have $\dot{x}(\cdot) \in L^\infty([1,\infty),\mathbf{R}^n)$, the limit $E_\infty := \lim_{t\to\infty} E(t)$ exists. Now we show $\dot{x}(\cdot) \in L^2([1,\infty),\mathbf{R}^n)$. To the end, define

$$E_1(t) = t^2(J(x(t)) - J(x^\dagger)) + \frac{1}{2}\left\|2(x(t) - x^\dagger) + t\dot{x}(t)\right\|^2 + 2(s-1)\left\|x(t) - x^\dagger\right\|^2$$

By using the convexity inequality $J(x^\dagger) \geqslant J(x) + (\nabla J(x), x^\dagger - x)$ for all $x \in \mathbf{R}^n$ and Eq.(19), it is not difficult to show that

$$\dot{E}_1(t) \leqslant -2(s-1)t\|\dot{x}(t)\|^2 \qquad (21)$$

Hence, for $s \geqslant 1$, $\dot{E}_1(t) \leqslant 0$ and thus $E_1(\cdot)$ is non-increasing. Together with the fact that $E_1(t) \geqslant 0$ for all $t \geqslant 0$, the limit $E_1(\infty) := \lim_{t\to\infty} E_1(t)$ exists.

Integrating both sides in Eq.(21), we obtain that

$$\int_1^\infty \left\|\dot{x}(t)\right\|^2 \mathrm{d}t \leqslant \int_1^\infty t\left\|\dot{x}(t)\right\|^2 \mathrm{d}t \leqslant \frac{E_1(1) - E_1(\infty)}{2(s-1)} < \infty \qquad (22)$$

which yields $\dot{x}(\cdot) \in L^2([1,\infty),\mathbf{R}^n)$. Combining it and $\dot{x}(\cdot) \in L^\infty([1,\infty),\mathbf{R}^n)$ to conclude that $\dot{x}(t) \to 0$ as $t \to \infty$.

(iii) From Eq.(19), we obtain

$$\ddot{x}(t) = -\frac{1+2s}{t}\dot{x}(t) + A^\mathrm{T}(b - Ax(t))$$

which gives immediately that $\ddot{x}(\cdot) \in L^\infty([1,\infty),\mathbf{R}^n)$.

(iv) Define

$$E_2(t) = t^2(J(x(t)) - J(x^\dagger)) + \frac{1}{2}\left\|2s(x(t) - x^\dagger) + t\dot{x}(t)\right\|^2 \qquad (23)$$

By using the convexity inequality $J(x^\dagger) \geqslant J(x) + (\nabla J(x), x^\dagger - x)$ and Eq.(19) again, we have

$$\dot{E}_2(t) \leqslant -2(s-1)t(J(x(t)) - J(x^\dagger))$$

Similar to $E_1(\cdot)$, for $s \geqslant 1$, $E_2(\cdot)$ is non-increasing, and the limit $\lim_{t\to\infty} E_2(t)$ exists. Therefore, from Eq.(23), there holds

$$t^2(J(x(t)) - J(x^\dagger)) \leqslant E_2(t) \leqslant E_2(1)$$

or

$$0 \leqslant J(x(t)) - J(x^\dagger) \leqslant E_2(1)t^{-2}$$

which gives $J(x(t)) - J(x^\dagger) = \mathcal{O}(t^{-2})$ as $t \to \infty$. The proof is completed.

Remark 1   Compared with the first order method (6), where the convergence rate of the objective functional is $J(x(t)) - J(x^\dagger) = \mathcal{O}(t^{-1})$, the convergence rates $J(x(t)) - J(x^\dagger) = o(t^{-1})$ in proposition 1 and $J(x(t)) - J(x^\dagger) = \mathcal{O}(t^{-2})$ in proposition 2 show that the second order dynamical systems (11) and (19) can achieve higher convergence order, indicating the second order dynamical system has a property of acceleration.

Now we are in a position to give the convergence result of the dynamical solution $x(t)$ of Eq.(7) with the noise-free data. Recall that $x^\dagger$ is the minimum norm minimizer of Eq.(4) with the exact data b. Then, formally, we have

$$x^\dagger = (A^\mathrm{T}A)^+ A^\mathrm{T}b$$

where the superscript " + " means the Moore-Penrose inverse.

Theorem 2   Let $x(\cdot)$ be the solution of Eq.(7) with the exact data, then

$$\lim_{t\to\infty} x(t) = x^\dagger \qquad (24)$$

Proof   Let $\{\sigma_j; u_j, v_j\}_{j=1}^\infty$ be the singular system for matrix A, i.e. we have $Av_j = \sigma_j u_j$ and $A^\mathrm{T}u_j = \sigma_j v_j$ with ordered singular values $\|A\| = \sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r > 0$.

Let's first consider the case with fixed damping parameter η. Similar to Ref.[32], we distinguish three different cases: (a) The overdamped case: $\eta > 2\|A\|$, (b) the underdamped case: there is an index j_0 such that $2\sigma_{j_0+1} < \eta < 2\sigma_{j_0}$, and (c) the critical damping case: an index j_0 exists such that $\eta = 2\sigma_{j_0}$. For simplicity, we only consider the overdamped case. The other two cases can be studied similarly. According to Ref.[32], we have

$$x(t) = (I - A^\mathrm{T}Ag(t,A^\mathrm{T}A))x_0 + \phi(t,A^\mathrm{T}A)\dot{x}_0 + g(t,A^\mathrm{T}A)A^\mathrm{T}b \qquad (25)$$

where

$$\begin{cases} g(t,\lambda) = \dfrac{1}{\lambda}\left(1 - \dfrac{\lambda_1}{\sqrt{\eta^2 - 4\lambda}}\mathrm{e}^{-\lambda_2 t} + \dfrac{\lambda_2}{\sqrt{\eta^2 - 4\lambda}}\mathrm{e}^{-\lambda_1 t}\right) \\ \phi(t,\lambda) = -\dfrac{1}{2\sqrt{\eta^2 - 4\lambda}}(\mathrm{e}^{-\lambda_2 t} - \mathrm{e}^{-\lambda_1 t}) \end{cases}$$

$$(26)$$

and $\lambda_1 = \dfrac{\eta + \sqrt{\eta^2 - 4\lambda}}{2}, \lambda_2 = \dfrac{\eta - \sqrt{\eta^2 - 4\lambda}}{2}$.

Together with theorem 1, it is straightforward to check that $x(t)$ defined Eq.(25) is a unique solution of Eq.(11). Moreover, we can deduce from Eq.(25) that $x(t) \to (A^{\mathrm{T}}A)^+ A^{\mathrm{T}}b$ as $t \to \infty$ by noting that

$$1 - \lambda g(t, \lambda) = \frac{\lambda_1}{\sqrt{\eta^2 - 4\lambda}} \mathrm{e}^{-\lambda_2 t} - \frac{\lambda_2}{\sqrt{\eta^2 - 4\lambda}} \mathrm{e}^{\lambda_1 t}$$

Next we consider the case with dynamic damping parameter $\eta(t) = (1 + 2s)/t$. According to Ref.[35], we have (note that $\dot{x}_0 = 0$)

$$x(t) = (I - A^{\mathrm{T}}A g_2(t, A^{\mathrm{T}}A))x_0 + g_2(t, A^{\mathrm{T}}A)A^{\mathrm{T}}b \tag{27}$$

where

$$g_2(t, \lambda) = \frac{1 - 2^s \Gamma(s + 1) \dfrac{J_s(\sqrt{\lambda} t)}{(\sqrt{\lambda} t)^s}}{\lambda} \tag{28}$$

$\Gamma(\cdot)$ and $J_s(\cdot)$ denote the Gamma function and the Bessel function of first kind of order s, respectively. As in the first case, $x(t)$ defined in Eq.(27) is the unique solution of Eq.(19). By using the asymptotic

$$J_s(\sqrt{\lambda} t) = \sqrt{\frac{2}{\pi}} (\sqrt{\lambda} t)^{-\frac{1}{2}} \cos\left(\sqrt{\lambda} t - \frac{\pi(2s + 1)}{4}\right) + \mathcal{O}\left(\frac{1}{\sqrt{\lambda} t}\right) \quad t \to \infty \tag{29}$$

for any fixed $\lambda > 0$, we conclude that $x(t) \to (A^{\mathrm{T}}A)^+ A^{\mathrm{T}}b$ as $t \to \infty$. The proof is completed.

1.3 Regularization properties for the noisy data

In section 1.2, we study the limiting behaviors of the solution for the second dynamical system (7) with noise-free data. It is shown that for exact data b, theoretically, the larger the time t is, the more accurate the solution $x(t)$ is. Therefore, we are readily given a satisfactory approximate solution $x(T^*)$ to $x^\dagger$ for a large enough time T^*. However, due to the inevitable noise, practically, it is not the case. In the case that only a noisy data b^δ is available, the system (7) cannot produce a reasonable approximate solution unless the terminating time T^* is chosen appropriately. In other word, the asymptotic process should be terminated in advance.

In this subsection, we investigate the regularization property of the dynamic solution $x^\delta(t)$ of system (7), equipped with the Morozov's discrepancy principle for selecting the terminating time T^*.

We define the Morozov's discrepancy function as

$$\chi(T) = \| Ax^\delta(T) - b^\delta \| - \tau\delta \tag{30}$$

where $x^\delta(t)$ is the solution of system (7), and τ a fixed positive number.

Proposition 3　For any $\tau > 0$, let the initial guess x_0 is chosen satisfying $\|Ax_0 - b^\delta\| \geqslant \tau\delta$. Then $\chi(T)$ has at least one zero point.

Proof　First, by theorem 1, it is easy to conclude that $\chi(\cdot)$ is continuous over $[t_0, \infty)$. Recall that x^δ is the minimum norm minimizer of Eq.(4). Then, formally, like $x^\dagger$ we have

$$x^\delta = (A^{\mathrm{T}}A)^+ A^{\mathrm{T}}b^\delta$$

Applying theorem 2 to $x^\delta(t)$ yield

$$\lim_{t \to \infty} x^\delta(t) = x^\delta$$

which implies

$$\lim_{T \to \infty} \chi(T) = \|Ax^\delta - b^\delta\| - \tau\delta = -\tau\delta < 0 \tag{31}$$

By combing Eq.(31) and $\chi(t_0) = \|Ax_0 - b^\delta\| - \tau\delta > 0$ as well as the continuity, we conclude that the nonlinear equation $\chi(T) = 0$ admits at least one root.

Now we are in a position to present a convergence result in the following.

Theorem 3　Let $x^\delta(t)$ be the solution of the dynamic system (7). Denote by $T^* = T(\delta, b^\delta)$ the first zero point of $\chi(T)$. Then $x^\delta(T^*)$ converges to $x^\dagger$ in $\mathbf{R}^n$ as $\delta \to 0$.

Proof　As $\delta \to 0$, $T(\delta, b^\delta)$ may go to infinity or may have a finite accumulation point. We first consider the situation that $T(\delta, b^\delta) \to \infty$. Recall that A is a $m \times n$ matrix with the rank r. We proceed to the proof with two possible cases: (i)$r \geqslant n$; (ii)$r < n$.

(i) In the case of $r \geqslant n$, $A^{\mathrm{T}}A$ is full rank, and thus

$$x^\dagger = (A^{\mathrm{T}}A)^{-1} A^{\mathrm{T}}b \quad x^\delta = (A^{\mathrm{T}}A)^{-1} A^{\mathrm{T}}b^\delta \tag{32}$$

By using triangle inequality,

$$\left\| x^\delta(T^*) - x^\dagger \right\| \leqslant \left\| x^\delta(T^*) - x^\delta \right\| + \left\| x^\delta - x^\dagger \right\| \tag{33}$$

On one hand, by applying the convergence in theorem 2, we have

$$\lim_{\delta \to 0} \left\| x^{\delta}(T^*) - x^{\delta} \right\| = 0 \qquad (34)$$

On the other hand, from Eq.(32)

$$\left\| x^{\delta} - x^{\dagger} \right\| \leqslant \left\| (AA^{\mathrm{T}})^{-1} \right\| \left\| A^{T}(b^{\delta} - b) \right\| \leqslant \frac{1}{\|A\|} \delta \qquad (35)$$

Combine Eqs.(33)—(35) to conclude that

$$\lim_{\delta \to 0} \left\| x^{\delta}(T^*) - x^{\dagger} \right\| = 0$$

(ii) For the case $r < n$, both the systems (1) and (3) have infinite solution, and $x^{\dagger}, x^{\delta}$ are the minimum norm solutions, respectively. Let $\{\delta_n\}$ be a sequence converging to 0 as $n \to \infty$, and let b^{δ_n} be a corresponding sequence of noisy data with $\left\| b^{\delta_n} - b \right\| \leqslant \delta_n$. For a pair of (δ_n, b^{δ_n}), denote by T_n^* the corresponding terminating time point determined from the discrepancy principles $\chi(T) = 0$. According to the continuity of $x^{\delta_n}(t)$, for any $\varepsilon > 0$ and large enough n and thus T_n^*, there exists a point $T < T_n^*$ such that

$$\left\| x^{\delta_n}(T_n^*) - x^{\delta_n}(T) \right\| \leqslant \varepsilon/3 \qquad (36)$$

From the continuity of $x^{\delta_n}(\cdot)$ due to theorem 2, for large enough n, there holds

$$\left\| x^{\delta_n}(T) - x(T) \right\| \leqslant \varepsilon/3 \qquad (37)$$

Moreover, by applying theorem 2 again, we obtain, for large enough T (because T_n^* could be large enough)

$$\left\| x(T) - x^{\dagger} \right\| \leqslant \varepsilon/3 \qquad (38)$$

Then by combing Eqs.(36)—(38) and using triangle inequality

$$\left\| x^{\delta_n}(T_n^*) - x^{\dagger} \right\| \leqslant \left\| x^{\delta_n}(T_n^*) - x^{\delta_n}(T) \right\| + \left\| x^{\delta_n}(T) - x(T) \right\| + \left\| x(T) - x^{\dagger} \right\|$$

There is a number n_1 such that for any $n \geqslant n_1$

$$\left\| x^{\delta_n}(T_n^*) - x^{\dagger} \right\| \leqslant \varepsilon$$

Since ε is arbitrary, we arrive at the convergence of $x^{\delta_n}(T_n^*)$ to $x^{\dagger}$ as $n \to \infty$.

For the case that $T(\delta, b^{\delta})$ has a finite accumulation point, we can use arguments similar to those in Theorem 2.4 of Ref.[36], and omit here. The proof is completed.

2　Iterative Schemes

For the numerical implementation, this section is devoted to present several iterative schemes for the resolution of problem (4). To the end, we first convert the system (7) into

$$\begin{cases} \dot{x}(t) = q(t) \\ \dot{q}(t) = -\eta(t)q(t) + A^{\mathrm{T}}(b^{\delta} - Ax(t)) \\ x(t_0) = x_0, \dot{x}(t_0) = \dot{x}_0 \end{cases} \quad (39)$$

or

$$\begin{cases} \dot{y}(t) = \begin{pmatrix} 0 & I \\ -A^{\mathrm{T}}A & -\eta(t)I \end{pmatrix}\begin{pmatrix} x(t) \\ q(t) \end{pmatrix} + \begin{pmatrix} 0 \\ A^{\mathrm{T}}b^{\delta} \end{pmatrix} \triangleq \\ \qquad A(t)y(t) + \bar{b} \\ x(t_0) = x_0, \dot{x}(t_0) = \dot{x}_0 \end{cases} \quad (40)$$

Then, the numerical discretization of the differential system (39) or (40) together with the discrepancy principle produces second order iterative regularization methods. The damped symplectic integrators are extremely attractive since these schemes are closely related to the canonical transformations[37], and the trajectories of the discretized second flow are usually more stable for its long-term performance. Applying the symplectic Euler method to the system (39), we obtain the iterative scheme at the kth step.

$$\begin{cases} q^{k+1} = q^k + \Delta t(A^{\mathrm{T}}(b^{\delta} - Ax^k) - \eta(t_k)q^k) \\ x^{k+1} = x^k + \Delta t q^{k+1} \\ x^0 = x_0, q^0 = \dot{x}_0 \end{cases} \quad (41)$$

where $x^k = x^{\delta}(t_k)$, $q^k = q(t_k)$, and Δt is the step-size. By elementary calculations, scheme (41) can be expressed as the form of following three-term semi-iterative method.

$$x^{k+1} = x^k + a_k(x^k - x^{k-1}) + \omega_k A^{\mathrm{T}}(b^{\delta} - Ax^k) \quad (42)$$

where $a_k = (1 - \Delta t \eta(t_k))$ and $\omega_k = (\Delta t)^2$.

For the high order Symplectic methods, the Störmer-Verlet scheme can be considered as

$$\begin{cases} q^{k+1/2} = q^k - \frac{\Delta t}{2}\eta(t_k)q^{k+1/2} + \frac{\Delta t}{2}A^{\mathrm{T}}(b^{\delta} - Ax^k) \\ x^{k+1} = x^k + \Delta t q^{k+1/2} \\ q^{k+1} = q^{k+1/2} - \frac{\Delta t}{2}\eta(t_{k+1})q^{k+1/2} + \frac{\Delta t}{2}A^{\mathrm{T}}(b^{\delta} - \\ \qquad Ax^{k+1}) \\ x^0 = x_0, q^0 = \dot{x}_0 \end{cases}$$

$$(43)$$

Like Eq.(41), the iteration (43) can also be rewritten in the form of Eq.(42), but with parameters

$$a_k = \frac{1 - \dfrac{\Delta t}{2}\eta(t_k)}{1 + \dfrac{\Delta t}{2}\eta(t_k)}, \omega_k = \frac{(\Delta t)^2}{1 + \dfrac{\Delta t}{2}\eta(t_k)}$$

In Ref.[35], a modified Störmer-Verlet scheme is proposed as

$$
\begin{cases}
q^{k+1/2} = q^k - \dfrac{\Delta t}{2}\eta(t_k)q^{k+1/2} + \dfrac{\Delta t}{2}A^{\mathrm{T}}(b^\delta - Ax^k) \\
x^{k+1} = x^k + \Delta t q^{k+1/2} \\
v^{k+1} = x^{k+1} + 2\Delta t a_{k+1}q^{k+1/2} \\
q^{k+1} = q^{k+1/2} - \dfrac{\Delta t}{2}\eta(t_{k+1})q^{k+1/2} + \\
\qquad \dfrac{\Delta t}{2}A^{\mathrm{T}}(b^\delta - Av^{k+1}) \\
x^0 = x_0, q^0 = \dot{x}_0
\end{cases} \tag{44}
$$

with $a_k = \dfrac{1 - \Delta t\eta(t_k)}{1 + \Delta t n(t_k)}$. The third step in Eq.(44) is inspired by the Nesterov's method. Unlike the Symplectic Euler method and Störmer-Verlet scheme, the scheme (44) expresses the recurrence form as

$$x^{k+1} = x^k + a_k(x^k - x^{k-1}) + \omega_k A^{\mathrm{T}}(b^\delta -$$
$$A(x^k + a_k(x^k - x^{k-1}))) \quad k = 1,2,\cdots \tag{45}$$

with parameters

$$a_k = \frac{1 - \dfrac{\Delta t}{2}n(t_k)}{1 + \dfrac{\Delta t}{2}\eta(t_k)}, \omega_k = \frac{(\Delta t)^2}{1 + \dfrac{\Delta t}{2}\eta(t_k)} \tag{46}$$

As indicated by theorems 2 and 3, theoretically, for the noise-free data or noisy data with small noisy level, the iteration should go far enough for a good enough approximate solution x^k. In this case, it requires long time behavior for the iteration. Keep in mind that we are solving ill-posed problems. Therefore, when the data contains noise with not small noisy level, the iteration should stop in advance before the solution gets worse. Therefore, the accuracy itself in the discretization of system (39) or (40) plays more important role than the requirement of the long behavior. We can apply high order Runge-Kutta methods for this purpose. As an example, the classical fourth order Runge-Kutta method is adopted as

$$
\begin{cases}
K_1 = A(t_k)y_k + \bar{b} \\
K_2 = A(t_{k+1/2})(y_k + \Delta t K_1/2) + \bar{b} \\
K_3 = A(t_{k+1/2})(y_k + \Delta t K_2/2) + \bar{b} \\
K_4 = A(t_{k+1})(y_k + \Delta t K_3) + \bar{b} \\
y_{k+1} = y_k + \dfrac{\Delta t}{6}(K_1 + 2K_2 + 2K_3 + K_4) \\
y_0 = (x_0, \dot{x}_0)^{\mathrm{T}}
\end{cases} \tag{47}
$$

We note that other Runge-Kutta methods are possible.

After the discretization, the regularization parameter is reduced to the iterative step k. For the schemes (41), (43), (44) and (47), the Morozov's discrepancy principle reads: find the smallest k satisfying

$$\|Ax_k - b^\delta\| \leqslant \tau\delta \tag{48}$$

3 Numerical Results

Two numerical examples are presented to demonstrate the effectiveness of the iterative regularization methods (41), (43), (44) and (47) for the resolution of the system (3). For the sake of simplicity, the symplectic Euler method (41) is termed as SE1 or SE2, corresponding to a constant or time-dependent damping parameter respectively. Similarly, the Störmer-Verlet method is termed as SV1 or SV2, the Modified Störmer-Verlet method is termed as MSV1 or MSV2, and the Runge-Kutta method is termed as RK1 or RK2. In the following experiments, the maximal iterative number $N_{\max}$ is set as 5 000.

For the comparison of the proposed methods with the existing work, we introduce such classical iteration methods as Landweber method, conjugate gradient method (termed as CG), ν-method and the Nesterov's method. For the problem (3), the Landweber method (5) reads

$$x_{k+1} = x_k - \Delta t A^{\mathrm{T}}(Ax_k - b^\delta) \tag{49}$$

with $0 < \Delta t < 2/\|A\|^2$. The ν-method has the form

$$x^{k+1} = x^k + \mu_{k+1}(x^k - x^{k-1}) - \omega_{\mathrm{norm}}\omega_k g_k \tag{50}$$

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

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

$$\omega_k = 4\frac{(2k+2\nu-1)(k+\nu-1)}{(k+2\nu-1)(2k+4\nu-1)}$$

where x^{-1}, x^0 are the initial values, and $g_k = A^T(Ax^k - b^\delta)$. In Eq.(50), ω_{norm} plays the role of normalization, and can be set as $1/\|A\|^2$.

The Nesterov's method is defined by

$$x^{k+1} = x^k + \frac{k-1}{k+\alpha-1}(x^k - x^{k-1}) - \omega g_k \quad (51)$$

with $\alpha > 3$, $0 < \omega \leqslant \omega_{norm}$ and $k = 0, 1, 2, \cdots$.

Like the methods in section 2, the Morozov's discrepancy principle (48) is applied to Landweber method, CG method, ν-method and the Nesterov method for the choice of iterative numbers.

3.1   Example 1

In the first example, a Fredholm integral equation of the first kind of convolution type in one space dimension is considered as[38-39]

$$(\mathcal{K}f)(x) = \int_0^1 k(x - x')f(x')\,d(x') = g(x)$$
$$0 < x < 1 \quad (52)$$

where the kernel function $k(x) = C\exp(-x^2/2\gamma^2)$ with positive constants C and γ. Using numerical integration, Eq.(52) is discretized to a linear system as

$$Kf = d \quad (53)$$

where $[K]_{ij} = hC\exp\left(-\dfrac{((i-j)h)^2}{2\gamma^2}\right)$, $[f]_i = f(ih)$, $[d]_i = g(ih)$, $h = 1/n$, $1 \leqslant i, j \leqslant n$. In the experiments of this subsection, fix $\gamma = 0.05$, $C = 1/\gamma$. The dependence of the condition number of the matrix K on the magnitude is plotted in Fig.1 which shows that K is ill-posed as long as n is slightly large. It is indicated in Fig.1 that the problem has the exponential ill-posedness.

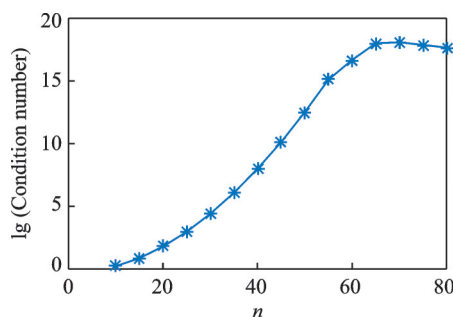For having the data in the right side of Eq.(53), we assume the exact solution $f^\dagger \equiv 1$.

Fig.1   Dependence of the condition number of the matrix K on n

Then $d = Kf^\dagger$, and the noisy data is constructed through

$$[d^\delta]_i = (1 + 2*(\text{rand}(1) - 0.5)*\delta')[d]_i$$
$$i = 1, \cdots, n$$

Then the noise level of the measurement data is calculated by $\delta = \|d^\delta - d\|$. With matrix K and data $d^\dagger$, the iterative methods (41), (43)—(44), and (47) are applied to solve the noisy system

$$Kf = d^\delta \quad (54)$$

In all methods, we simply set the initial values $f_0 = 0$, $\dot{f}_0 = 0$ and the magnitude $n = 100$. For an approximate solution f^k, its accuracy is assessed by using the relative error as

$$\text{L2err} = \|f^k - f^\dagger\|/\|f^\dagger\|$$

In addition, IterN is used to denote iterative number where the iteration stops.

The effect of the step-size Δt on the iterative number and the solution accuracy is firstly investigated. To the end, fix $\delta' = 1\%$, $\tau = 1.03$. The results are shown in Figs.2 and 3. Fig.2 corresponds to constant damping parameters ($\eta = 0.6$ for SE1, $\eta = 0.8$ for SV1, $\eta = 0.1$ for MSV1 and $\eta = 0.1$ for RK1) while Fig.3 corresponds to a variable damping parameter ($\eta(t) = 4/t$ for all methods). It is
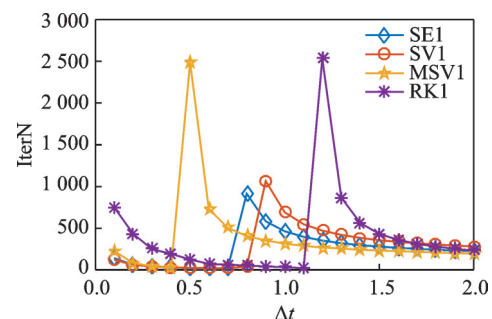
(a) Dependence of IterN on Δt
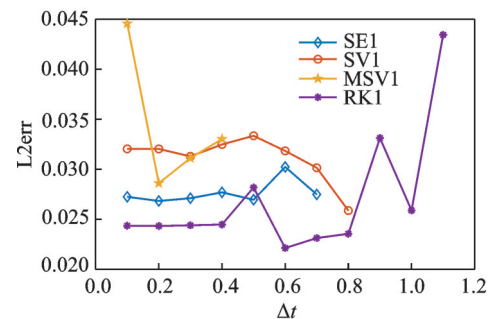
(b) Dependence of L2err on Δt

Fig.2   Dependence of IterN and L2err on Δt with η as a constant
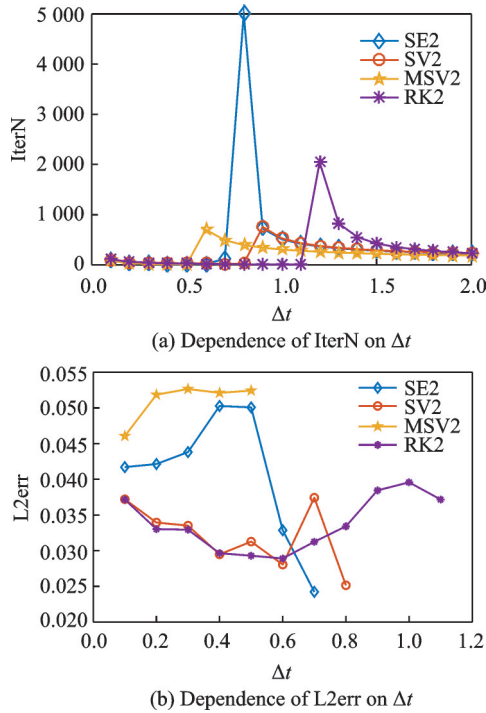
(a) Dependence of IterN on Δt

(b) Dependence of L2err on Δt

Fig.3　Dependence of IterN and L2err on Δt with $\eta(t) = 4/t$

(a) Dependence of IterN on τ

(b) Dependence of L2err on τ

Fig.4　Dependence of IterN and L2err on τ with η as a constant

(a) Dependence of IterN on τ

(b) Dependence of L2err on τ

Fig.5　Dependence of IterN and L2err on τ with $\eta(t) = 4/t$

concluded that for all methods, on one hand, the larger Δt is, the faster the iteration is; on the other hand, its too big values may lead to the divergence. In the following experiments, we set $\Delta t = 0.7$ for SE1, $\Delta t = 0.8$ for SV1, $\Delta t = 0.4$ for MSV1, $\Delta t = 1.1$ for RK1, $\Delta t = 0.6$ for SE2, $\Delta t = 0.8$ for SV2, $\Delta t = 0.4$ for MSV2, $\Delta t = 1.1$ for RK2. They are all approximately optimal.

Next we investigate the effect of the parameter τ, used in discrepancy principle, on the iterative number and the solution accuracy. For this purpose, fix $\delta' = 1\%$. The results are plotted in Figs. 4 and 5 which show that on the whole, the larger the value of τ is, the less the iterative number is and the worse the solution accuracy is. In the remaining part, without specific statement, we set $\tau = 1.03$.

We further compare the iterative schemes with the existing Landweber method, CG method, ν-method and Nesterov method. We set $\Delta t = 0.3$ for Landweber method, $\alpha = 3, \omega = 0.16$ for Nesterov method, $\Delta t = 0.7, \eta = 0.6$ for SE1, $\Delta t = 0.8, \eta = 0.8$ for SV1, $\Delta t = 0.4, \eta = 0.1$ for MSV1, $\Delta t = 1.1, \eta = 0.1$ for RK1, $\Delta t = 0.6$ for SE2, $\Delta t = 0.8$ for SV2, $\Delta t = 0.4$ for MSV2, $\Delta t = 1.1$ for RK2.
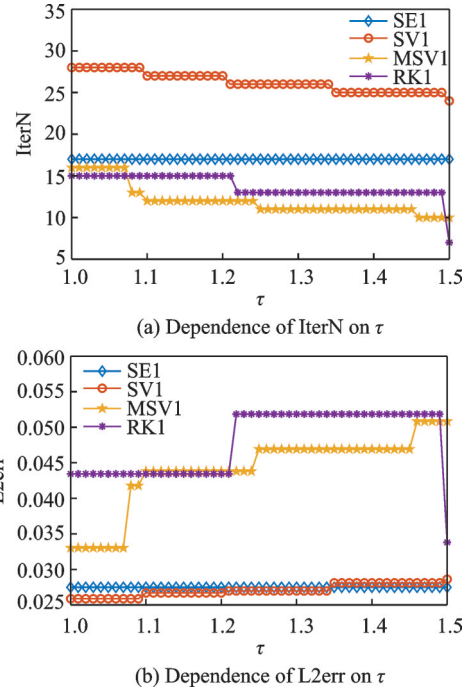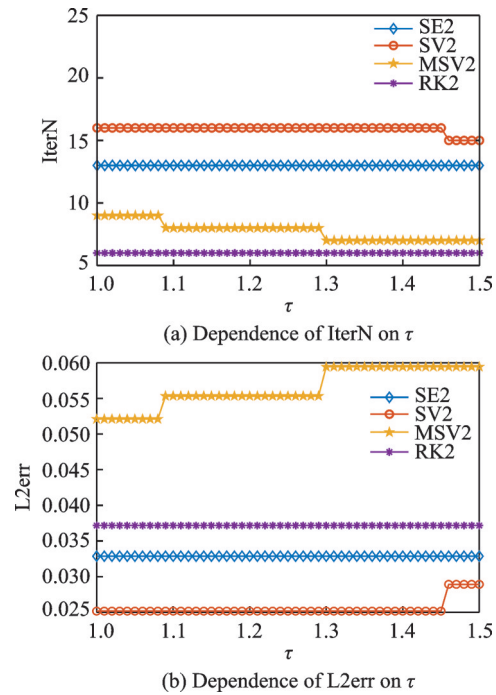
Moreover, set $\tau = 1.03$ in all methods. The iterative numbers and relative errors in approximate solutions for different methods and different noise levels are given in Table 1, from which we conclude that with properly chosen parameters, all the mentioned methods are stable and can produce satisfactory solu-

Table 1　Comparison between different methods in Example 1

Methods	$\delta'/\%$					
	0.1		1		5	
	L2err	IterN	L2err	IterN	L2err	IterN
Landweber	2.210 2e−2	112	3.597 4e−2	28	6.643 8e−2	19
CG	1.899 0e−2	7	4.484 5e−2	3	9.707 2e−2	2
$\nu = 0.5$	2.260 6e−2	541	2.347 9e−2	55	7.699 8e−2	12
$\nu = 0.7$	9.654 9e−3	99	2.606 0e−2	19	7.356 5e−2	6
$\nu = 1.0$	1.775 8e−2	33	3.781 7e−2	9	7.055 8e−2	4
$\nu = 1.5$	2.214 5e−2	25	4.987 9e−2	6	7.367 7e−2	3
$\nu = 2.0$	2.224 0e−2	27	5.193 3e−2	6	7.418 5e−2	3
Nesterov	2.020 8e−2	44	4.996 1e−2	9	8.645 1e−2	3
SE1	2.108 3e−2	34	2.747 6e−2	17	6.740 9e−2	16
SV1	1.993 7e−2	49	2.586 3e−2	28	6.963 8e−2	16
MSV1	1.578 8e−2	52	3.302 0e−2	16	8.175 9e−2	3
RK1	1.047 2e−2	49	4.343 1e−2	15	8.028 2e−2	7
SE2	1.459 1e−2	56	3.285 9e−2	13	6.363 5e−2	8
SV2	1.303 0e−2	53	2.515 2e−2	16	6.978 8e−2	14
MSV2	2.212 1e−2	40	5.211 2e−2	9	8.725 2e−2	4
RK2	2.073 2e−2	16	3.717 0e−2	6	6.451 0e−2	5

tions. Compared with the Landweber method, all the other mentioned methods have acceleration effect. On the whole, for all methods, the larger the noise level is, the worse the solution accuracy is, but the less the required iterative number is. Compared with SE1, SV1, MSV1 and RK1 methods, SE2, SV2, MSV2 and RK2 have better behavior, such as fewer iterative steps or better solution accuracy. Particularly, Table 1 shows that RK2 possesses the best behavior in both the iterative number and solution accuracy.

Finally we study the effect of the magnitude of the problem on different iterative schemes. We fix $\delta' = 1\%$, $\tau = 1.03$. The experiments for Landweber method, ν-method, Nesterov method and RK2 are implemented repeatedly to solve Eq.(54) with $n = 25$, 50, 100, 200, 400, 800, 1 600 and 3 200, respectively, and the results are shown in Fig.6. It is indicated from the Fig.6 that when the order of the matrix increases, the accuracy of the solution and the number of iteration steps change little. Moreover, Fig.6 also makes it clear that the RK2 performs better than three existing methods, especially when the scale of the problem increases.

(a) Dependence of IterN on the scale n

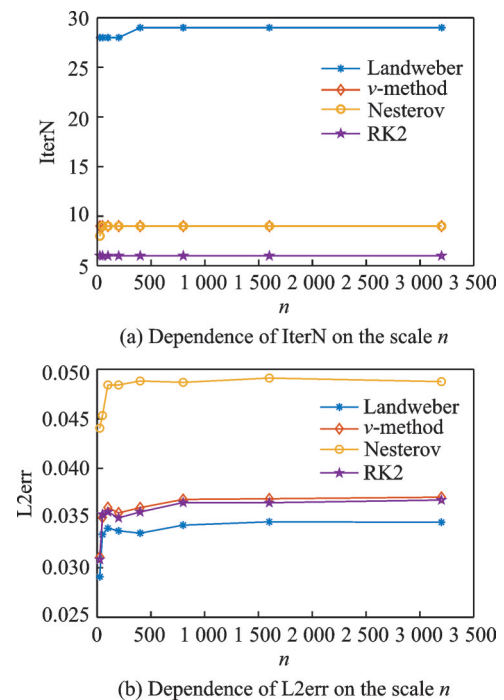
(b) Dependence of L2err on the scale n

Fig.6　Dependence of IterN and L2err on the scale n

3.2　Example 2

In the second example, the well-known ill-posed Hilbert matrix is taken to model the operator as

$$A = \begin{pmatrix} 1 & \dfrac{1}{2} & \dfrac{1}{3} & \cdots & \dfrac{1}{n} \\ \dfrac{1}{2} & \dfrac{1}{3} & \dfrac{1}{4} & \cdots & \dfrac{1}{n+1} \\ \cdots & \cdots & \cdots & \cdots & \\ \dfrac{1}{n} & \dfrac{1}{n+1} & \dfrac{1}{n+2} & \cdots & \dfrac{1}{2n-1} \end{pmatrix}$$

The dependence of the condition number of the matrix A on its scale is plotted in Fig.7, which shows that the condition number of A increases dramatically and thus A is ill-posed. Morevoer, like Example 1, assume again the exact solution $x^{\dagger} = (1, 1, \cdots, 1)^{\mathrm{T}}$ and compute $Ax^{\dagger}$ for the right side b. The noisy data b^{δ} is constructed like Example 1. Then iterations (41), (43)—(44) and (47) are applied to solve
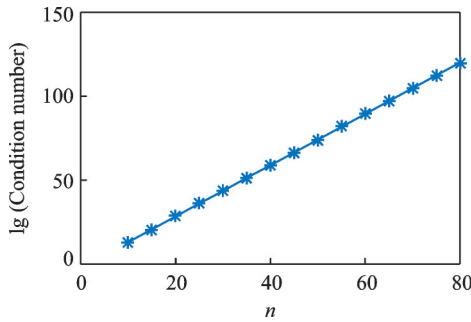
$$Ax = b^{\delta} \tag{55}$$

Fig.7   Dependence of condition number of A on the scale n

The dependence of the iterative numbers and the accuracy in approximate solutions obtained with iterations in Section 3 on parameters Δt and τ are plotted in Figs. 8—11. Specifically, Fig. 8 is fo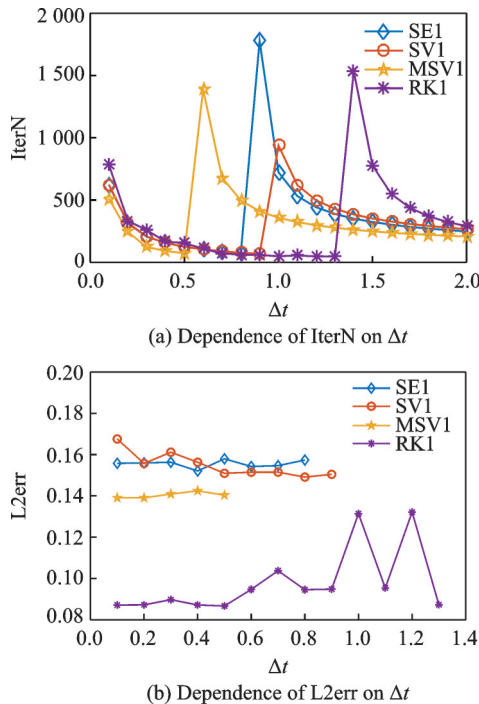r constant η (with $\eta = 0.2$ for SE1, $\eta = 0.2$ for SV1, $\eta = 0.1$ for MSV1, $\eta = 0.1$ for RK1) while Fig.9 is for $\eta(t) = 4/t$ (both with $\tau = 1.03$); Fig.10 is for constant η (with $\Delta t = 0.8$, $\eta = 0.2$ for SE1, $\Delta t = 0.9$, $\eta = 0.2$ for SV1, $\Delta t = 0.5, \eta = 0.1$ for MSV1, $\Delta t = 1.2$, $\eta = 0.1$ for RK1) while Fig.11 is for $\eta(t) = 4/t$ (with $\Delta t = $
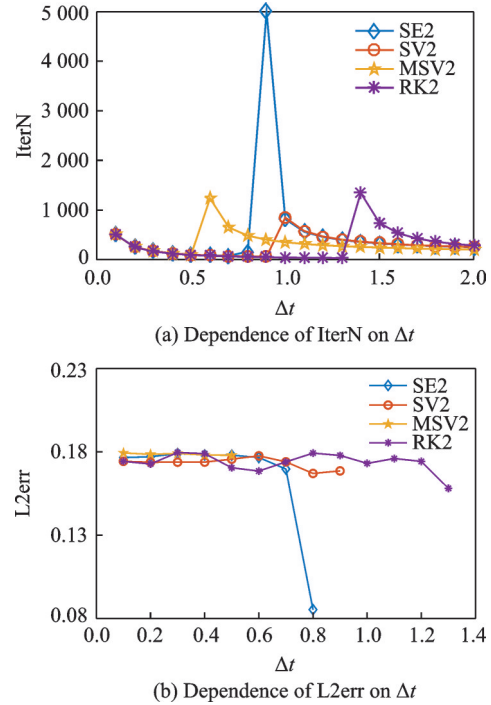

(a) Dependence of IterN on Δt

(b) Dependence of L2err on Δt

Fig.9   Dependence of IterN and L2err on Δt with $\eta(t) = 4/t$

(a) Dependence of IterN on Δt

(b) Dependence of L2err on Δt

Fig.8   Dependence of IterN and L2err on Δt with η as a constant

(a) Dependence of IterN on τ

(b) Dependence of L2err on τ

Fig.10   Dependence of IterN and L2err on τ with η as a constant
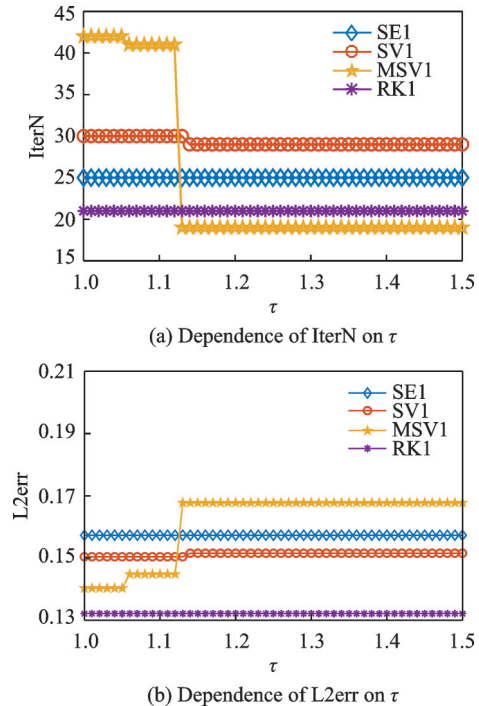
(a) Dependence of IterN on τ

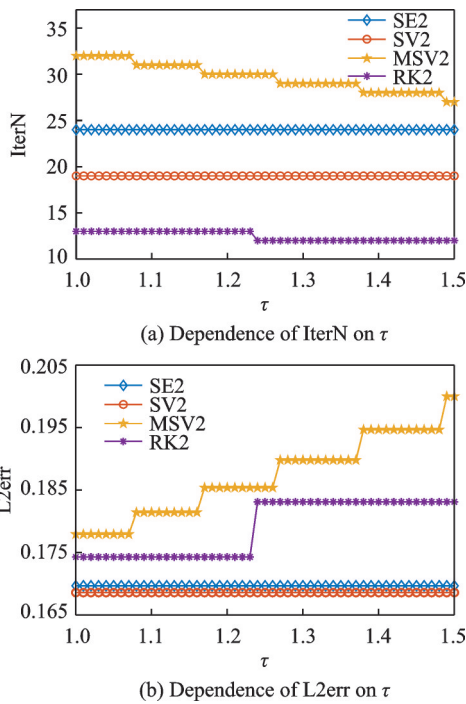(b) Dependence of L2err on τ

Fig.11　Dependence of IterN and L2err on τ with $\eta(t) = 4/t$

0.7 for SE2, $\Delta t = 0.9$ for SV2, $\Delta t = 0.5$ for MSV2, $\Delta t = 1.1$ for RK2). In all figures, $\delta' = 1\%$. Similar conclusion to those for Example 1 can be drawn from these figures.

Similar to those in Example 1, a comparison between different methods for different noise data is displayed in Table 2, where $\Delta t = 0.3$ for Landweber method, $\alpha = 3, \omega = 0.2$ for Nesterov method, $\Delta t = 0.8, \eta = 0.2$ for SE1, $\Delta t = 0.9, \eta = 0.2$ for SV1, $\Delta t = 0.5, \eta = 0.1$ for MSV1, $\Delta t = 1.2, \eta = 0.1$ for RK1, $\Delta t = 0.7$ for SE2, $\Delta t = 0.9$ for SV2, $\Delta t = 0.5$ for MSV2, $\Delta t = 1.1$ for RK2. Also, $\tau = 1.03$ in all methods. It can be concluded from Table 2 that with properly chosen parameters, all the mentioned methods are stable and can produce satisfactory solutions.

Table 2　Comparison between different methods in Example 2

Methods	$\delta'/\%$					
	0.1		1		5	
	L2err	IterN	L2err	IterN	L2err	IterN
Landweber	8.124 2e−2	2461	1.766 5e−1	126	3.450 7e−1	11
CG	6.941 6e−2	4	1.389 1e−1	3	2.918 4e−1	2
$\nu = 0.5$	2.953 9e−2	545	8.479 6e−2	60	2.065 4e−1	12
$\nu = 0.7$	7.475 3e−2	130	1.476 7e−1	28	2.946 0e−1	8
$\nu = 1.0$	8.104 4e−2	82	1.758 0e−1	18	3.472 2e−1	5
$\nu = 1.5$	8.171 0e−2	98	1.789 2e−1	21	3.434 2e−1	6
$\nu = 2.0$	8.171 2e−2	113	1.750 0e−1	25	3.488 0e−1	6
Nesterov	7.866 6e−2	159	1.755 5e−1	32	3.494 9e−1	8
SE1	8.138 8e−2	180	1.440 2e−1	29	1.275 0e−1	29
SV1	8.141 4e−2	160	1.206 3e−1	46	1.330 7e−1	28
MSV1	8.190 3e−2	136	1.403 3e−1	42	2.669 2e−1	11
RK1	7.916 7e−2	66	1.321 6e−1	21	2.354 1e−1	5
SE2	7.870 6e−2	114	1.696 6e−1	24	2.541 0e−1	14
SV2	7.949 4e−2	86	1.685 7e−1	19	2.704 6e−1	10
MSV2	8.159 0e−2	148	1.779 1e−1	32	3.513 1e−1	8
RK2	8.121 3e−2	67	1.760 0e−1	14	3.116 8e−1	6

The effect of the problem magnitude on different iterative schemes is also studied. The experiments for Landweber method, ν-method, Nesterov method and RK2 are implemented repeatedly to solve Eq. (55) with $\delta' = 1\%, \tau = 1.03$, and $n = 25, 50, 100, 200, 400, 800, 1\ 600$ and $3\ 200$, respectively, whose results are shown in Fig.12. Again we can see from Fig.12 that the RK2 performs the best when compared with the existing classical methods.

(a) Dependence of IterN on the scale n

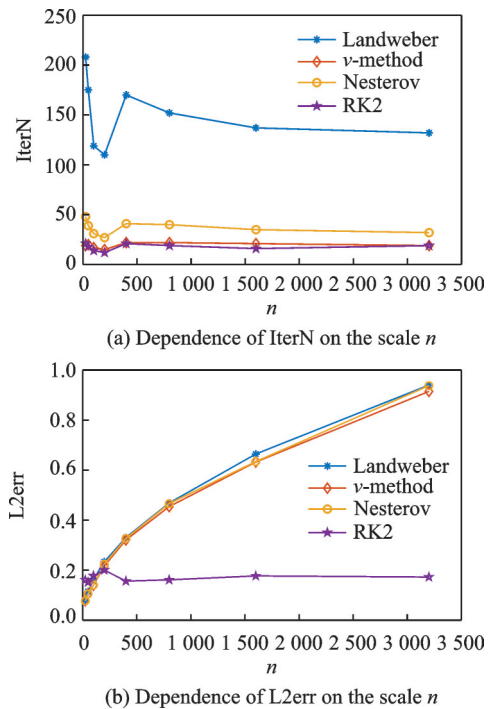(b) Dependence of L2err on the scale n

Fig.12 Dependence of IterN and L2err on the scale n

4 Conclusions

The newly developed second order asymptotic system is applied to solve severely ill-posed linear system. Compared with the existing work, a series of theoretical results are presented and proved with simplified arguments. Since the existing work for second order system often assumes that it has exact solution while our problem may have no solution, the analysis is adjusted. On the whole, the fourth order Runge-Kutta method is better than the classical ν-method and Nesterov method, and much faster than Landweber method. Moreover, it is more obvious when the scale and the ill-posedness of the system increase. The reason is that the system is ill-posed and iteration should stop in advance before the accuracy gets worse. Therefore, although it is not a symplectic method, it has high precision. This is a new sight. Better behavior can be expected if higher precision scheme is used. The idea can also be applied to other linear inverse problems such as inverse source problems and Cauchy problems etc., which will be studied in the future.

References

[1]   BERNTSSON F, ELDÉN L. Numerical solution of a Cauchy problem for the Laplace equation[J]. Inverse Problems, 2001, 17(4): 839-853.

[2]   ANDRIEUX S, ABDA A B, BARANGER T N. Data completion via an energy error functional[J]. Comptes Rendus-Mécanique, 2005, 333(2): 171-177.

[3]   ANDRIEUX S, BARANGER T N. An energy error-based method for the resolution of the Cauchy problem in 3D linear elasticity[J]. Computer Methods in Applied Mechanics and Engineering, 2008, 197(9/10/11/12): 902-920.

[4]   ISAKOV V. Inverse problems for partial differential equations[M]. 2nd Edition. New York: Springer, 2006.

[5]   LIU Y, ZHANG X, LU M W. A meshless method based on least-squares approach for steady- and unsteady-state heat conduction problems[J]. Numerical Heat Transfer, Part B (Fundamentals), 2005, 47(3): 257-275.

[6]   MARTIN T J, DULIKRAVICH G S. Inverse determination of boundary conditions and sources in steady heat conduction with heat generation[J]. Heat Transfer, 1996, 118(3): 546-554.

[7]   LIU X D, ZHANG B, YANG J Q. The inverse electromagnetic scattering problem in a piecewise homogeneous medium[J]. Inverse Problems, 2010, 26(12): 125001.

[8]   ZHENG J C, CHENG J, LI P, et al. Periodic surface identification with phase or phaseless near-field data[J]. Inverse Problems, 2017, 33(11): 115004.

[9]   HU G H, LI L, ZOU J. Unique determination of a penetrable scatterer of rectangular type for inverse Maxwell equations by a single incoming wave[J]. Inverse Problems, 2019, 35(3): 035006.

[10] ZHENG W, LI S, KROL A, et al. Sparsity promoting regularization for effective noise suppression in SPECT image reconstruction[J]. Inverse Problems, 2019, 35(11): 115011.

[11] CHEN P J, HUANG J G, ZHANG X Q. A primal-dual fixed point algorithm for convex separable minimization with applications to image restoration[J]. Inverse Problems, 2013, 29(2): 025011.

[12] JU Rong, QIU Xiaohui, ZHAO Yang, et al. Fast reconstruction of radiated noise source and mechanism description method[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2012, 44(2): 217-221. (in Chinese)

[13] GOLUB G H, VAN LOAN C F. Matrix computations[M]. 3rd Edition. Baltimore: The John Hopkings University Press, 1996.

[14] KAIPIO J, SOMERSALO E. Statistical and compu-

tational inverse problems[M]. New York: Springer, 2005.

[15] SHEA J D, VAN VEEN B D, HAGNESS S C. A TSVD analysis of microwave inverse scattering for breast imaging[J]. IEEE Transactions on Biomedical Engineering, 2012, 59(4): 936-945.

[16] CHAVENT G, KUNISCH K. Convergence of Tikhonov regularization for constrained ill-posed inverse problems[J]. Inverse Problems, 1994, 10(1): 63-76.

[17] ENGL H W, KUNISCH K, NEUBAUER A. Convergence rates for Tikhonov regularization of non-linear ill-posed problems[J]. Inverse Problems, 1989, 5(4): 523-540.

[18] ALIFANOV O M, RUMYANTSEV S V. Application of iterative regularization for the solution of incorrect inverse problems[J]. Journal of Engineering Physics & Thermophysics, 1987, 53(5): 1335-1342.

[19] BAKUSHINSKY A B, KOKURIN M M, KOKURIN M Y. Regularization algorithms for ill-posed problems[M]. Berlin: De Gruyter, 2018.

[20] ALBER Y, RYAZANTSEVA I. Nonlinear ill-posed problems of monotone type[M]. Dordrecht: Springer, 2006.

[21] BAKUSHINSKY A B, KOKURIN M Y. Iterative methods for approximate solution of inverse problems[M]. Dordrecht: Springer, 2004.

[22] DOICU A, TRAUTMANN T, SCHREIER F. Numerical regularization for atmospheric inverse problems[M]. Berlin: Springer, 2010.

[23] WANG Z S, MU L M, LIU R R, et al. Arnoldi projection fractional Tikhonov for large scale ill-posed problems[J]. Transactions of Nanjing University of Aeronautics & Astronautics, 2018, 35(3): 395-402.

[24] YANG W Q, SPINK D M, YORK T A, et al. An image-reconstruction algorithm based on Landweber's iteration method for electrical-capacitance tomography[J]. Measurement Science and Technology, 1999, 10(11): 1065-1069.

[25] KALTENBACHER B, NEUBAUER A, SCHERZER O. Iterative regularization methods for nonlinear ill-posed problems[M]. Berlin: Walter de Gruyter GmbH & Co. KG, 2008.

[26] TAUTENHAHN U. On the asymptotical regularization of nonlinear ill-posed problems[J]. Inverse Problems, 1994, 10(6): 1405-1418.

[27] ZHANG Y, GONG R F. Second order asymptotical regularization methods for inverse problems in partial differential equations[J]. Journal of Computational and Applied Mathematics, 2020, 375: 112798.

[28] ENGL H W, HANKE M, NEUBAUER A. Regularization of inverse problems[M]. Dordrecht: Kluwer Academic Publishers, 1996.

[29] JIN Q N. On a regularized Levenberg-Marquardt method for solving nonlinear inverse problems[J]. Numerische Mathematik, 2010, 115(2): 229-259.

[30] JIN Q N, TAUTENHAHN U. On the discrepancy principle for some Newton type methods for solving nonlinear inverse problems[J]. Numerische Mathematik, 2009, 111(4): 509-558.

[31] NEUBAUER A. On Nesterov acceleration for Landweber iteration of linear ill-posed problems[J]. Journal of Inverse and Ill-Posed Problems, 2017, 25(3): 381-390.

[32] ZHANG Y, HOFMANN B. On the second order asymptotical regularization of linear ill-posed inverse problems[J]. Applicable Analysis, 2020, 99(6): 1000-1025.

[33] BREZIS H. Functional analysis, Sobolev spaces and partial differential equations[M]. New York: Springer, 2011.

[34] FENG Z H, LI F Y, LYU Y, et al. A note on Cauchy-Lipschitz-Picard theorem[J]. Journal of Inequalities and Applications, 2016(1): 271.

[35] GONG R F, HOFMANN B, ZHANG Y. A new class of accelerated regularization methods, with application to bioluminescence tomography[J]. Inverse Problems, 2020, 36(5): 055013.

[36] HANKE M, NEUBAUER A, SCHERZER O. A convergence analysis of the Landweber iteration for nonlinear ill-posed problems[J]. Numerische Mathematik, 1995, 72(1): 21-37.

[37] HAIRER E, LUBICH C, WANNER G. Geometric numerical integration: Structure-preserving algorithms for ordinary differential equations[M]. 2nd ed. Berlin: Springer-Verlag, 2006.

[38] VOGEL C R. Computational methods for inverse problems[M]. Philadelphia: SIAM, 2002.

[39] ZHANG Y, LUKYANENKO D, YAGOLA A. An optimal regularization method for convolution equations on the sourcewise represented set[J]. Journal of Inverse and Ill-Posed Problems, 2015, 23(5): 465-475.

Authors　Dr. **GONG Rongfang** is currently an associate professor in Department of Mathematics，Nanjing University of Aeronautics and Astronautics. His research interests include computational methods for inverse problems and optimal control，bioluminescence tomography and parameter estimation.

Ms. **HUANG Qin** is currently a postgraduate student of computational mathematics in Department of Mathematics，Nanjing University of Aeronautics and Astronautics. Her research interest focuses on inverse problem.

（Production Editor：WANG Jing）

求解严重病态线性方程组的基于时间离散的迭代正则化方法

龚荣芳，黄　沁

（南京航空航天大学理学院,南京 211106,中国）

摘要：近年来,反问题在计算数学中得到了越来越多的关注,在工程应用中也越来越重要。许多反问题在离散化后都退化为线性方程组。由于反问题的典型病态性,退化的线性方程组通常也是病态的,特别是当其规模很大时,这就给计算带来了很大的困难。特别地,病态线性方程组右端的一个小扰动可能会引起解的显著变化。因此需要采用正则化方法来获得稳定解。本文应用一类新的加速迭代正则化方法来求解这类大规模不适定线性方程组。一个迭代格式只有在迭代提前终止时才称为正则化方法。本文采用 Morozov 偏差原理作为终止准则。与传统的 Landweber 迭代法相比,新方法具有加速效果,可以与著名的 ν 方法和 Nesterov 方法相媲美。从数值结果可以看出,采用适当的离散化格式,本文的方法甚至比 ν 方法和 Nesterov 方法有更好的行为。

关键词：线性方程组;不适定性;大规模;迭代正则化方法;加速