Computational Methods in Applied Mathematics
Vol. 13 (2013), No. 3, pp. 349–362

Doi: 10.1515/cmam-2013-0009

DE GRUYTER

A Fast Algorithm for Parameter Identification Problems Based on the Multilevel Augmentation Method

Hui Cao · M. Thamban Nair

Abstract — A multilevel augmentation method is considered to solve parameter identification problems in elliptic systems. With the help of the natural linearization technique, the identification problems can be transformed into a linear ill-posed operation equation, where noise exists not only in RHS data but also in operators. Based on multiscale decomposition in solution space, the multilevel augmentation method leads to a fast algorithm for solving discretized ill-posed problems. Combining with Tikhonov regularization, in the implementation of the multilevel augmentation method, one only needs to invert the same matrix with a relatively small size and perform a matrix-vector multiplication at the linear computational complexity. As a result, the computation cost is dramatically reduced. The a posteriori regularization parameter choice rule and the convergence rate for the regularized solution are also studied in this work. Numerical tests illustrate the proposed algorithm and the theoretical estimates.

2010 Mathematical subject classification: 35R30, 65J15, 65J22.

Keywords: Parameter Identification, Multilevel Augmentation Method, Balancing Principle.

1. Introduction

Parameter identification problems in PDEs are typical inverse problems which are ill-posed, as opposed to the forward problems of solving PDEs. In this paper, we consider the problem of identifying a distributed parameter $a = a(x)$ from noisy measurements u^δ of the solution of the boundary value problem

$$\begin{aligned} -\nabla(a\nabla u) &= f \quad \text{in } \Omega, \\ u &= g \quad \text{on } \partial\Omega. \end{aligned} \tag{1.1}$$

Here Ω is a convex domain with Lipschitz boundary in the Euclidean space $\mathbb{R}^d$, $f \in L^2(\Omega)$, and $g \in H^{\frac{3}{2}}(\Omega)$, and for some fixed noise level δ, we have

$$\|u^\delta - u\|_{L_2} \leqslant \delta.$$

The inverse problem above plays an important role in many scientific and industrial applications. For example, it can model the inverse ground water filtration problem of reconstruction

Hui Cao
School of Mathematics and Computational Science, Sun Yat-sen University, Guangzhou 510275, P. R. China
E-mail: `caohui6@mail.sysu.edu.cn`.

M. Thamban Nair
Department of Mathematics, IIT Madras, Chennai 600036, India
E-mail: `mtnair@iitm.ac.in`.

of the diffusivity a of a sediment from the measurements of the piezometric head u in the steady state case.

Here we assume the existence of the exact solution $a^\dagger$ corresponding to the unperturbed data u. On the other hand, if a is known on the boundary of Ω and $\triangle u$ is bounded away from zero, then the uniqueness of a on Ω can be also proved (see, e.g., [2]). In this paper, we are interested in the numerical reconstruction of the parameter in the considered problem.

Although (1.1) is a linear elliptic equation, the relation between a and u is obviously nonlinear. Meanwhile, the inverse problem of identifying a from the noisy measurements of u is also ill-posed in the sense that arbitrarily small perturbation on u can lead to arbitrarily large deviation in the solution a in any meaningful topology. A major cause of such an instability is that in the process of the identification problem, data differentiation is usually unavoidable. Therefore, a suitable regularization method has to be applied to obtain a reconstruction with desired accuracy.

The parameter identification problem (1.1) is usually treated as a nonlinear operator equation

$$F(a) = u, \tag{1.2}$$

where F is a nonlinear "coefficient-to-solution" mapping, which, for example can be considered as acting from an appropriate subset of $L^\infty(\Omega)$ to $H^1(\Omega)$. For solving a nonlinear ill-posed problem of the form (1.2), iterative regularization methods, such as Landweber iteration or Newton-type methods, are usually applied (see, e.g., [15, 21]). The idea of these iterative methods consist of linearizing repeatedly the nonlinear equation (1.2) around some approximate solution obtained from the previous iterations. However, since the adjoint of the Fréchet derivative of the nonlinear forward mapping must be calculated many times, the above methods require a lot of numerical efforts. In [15], a linearization technique was proposed, and it has been effectively applied in [3]. Such a linearization method is essentially different from iterative regularization methods. It makes use of the structure of the elliptic PDE and avoids the appearance of the adjoint operator which is usually complicated for computational purposes. As in [3], we call the above mentioned linearization procedure the *natural linearization* (NL) technique.

In this paper, for a linearized form of the parameter identification problem (1.1), we develop a fast algorithm based on the so-called *multilevel augmentation method* (MAM) which was first proposed in [8] for solving operator equations. Later the MAM algorithm has been further investigated in [10, 11] for solving linear integral equations and ordinary differential equations. The problems considered in the literatures mentioned above are linear and well-posed. In [12, 13], MAM is further utilized for solving ill-posed operator equations combined with the Galerkin method and the collocation method, respectively. MAM is an iterative algorithm based on the multiscale structure of the discretized linear system. The main feature of the MAM algorithm is that during the process of resolution, one only needs to invert the same matrix with a relatively small size and perform a matrix-vector multiplication at the linear computational complexity. In this sense, the computation cost is dramatically reduced. We will apply the MAM algorithm combined with Tikhonov regularization and solve the linear ill-posed operator equation obtained by the NL technique, where both the operator and the data on the right-hand side are noisy. Moreover, for the choice of the regularization parameter, we shall use the recently developed a posteriori parameter choice rule based on the balancing principle (cf., e.g., [5, 19, 20]). By the balancing principle, the regularization parameter is chosen corresponding to a simulated solution with order-optimal accuracy. At the same time it can provide an estimate to the constant in stability bound which indicates some accuracy of noise level in numerics. In the last section of the paper, we present some numerical tests data to illustrate the algorithm as well as the theoretical estimates.

2. Natural Linearization and Multilevel Augmentation Method

2.1. Reformulation of the Problem by Natural Linearization

As in [3,15], using an initial guess a_0, we can represent (1.1) as follows:

$$\begin{aligned} -\nabla(a_0\nabla(u-u_0)) &= \nabla((a-a_0)\nabla u) \quad &&\text{in } \Omega, \\ u-u_0 &= 0 &&\text{on } \partial\Omega, \end{aligned}$$

where u_0 solves

$$\begin{aligned} -\nabla(a_0\nabla u_0) &= f \quad &&\text{in } \Omega, \\ u_0 &= g &&\text{on } \partial\Omega. \end{aligned}$$

Then a linear operator equation can be obtained:

$$\bar{A}s = \bar{r}, \tag{2.1}$$

where $s = a - a_0$ is the difference between the unknown parameter a and the initial guess a_0, $\bar{r} = u - u_0$, and the operator $\bar{A}$ maps s to the solution z of

$$\begin{aligned} -\nabla(a_0\nabla z) &= \nabla(s\nabla u) \quad &&\text{in } \Omega, \\ z &= 0 &&\text{on } \partial\Omega. \end{aligned} \tag{2.2}$$

Replacing u by a smoothed version u^δ_{sm} of u^δ such that $\nabla(s\nabla u^\delta_{\text{sm}})$ is well-defined, we switch to the equation

$$As = r^\delta, \tag{2.3}$$

with perturbed operator $A = A(u^\delta)$ and noisy right-hand side $r^\delta = u^\delta_{\text{sm}} - u_0$, where A maps s to the solution z of

$$\begin{aligned} -\nabla(a_0\nabla z) &= \nabla(s\nabla u^\delta_{\text{sm}}) \quad &&\text{in } \Omega, \\ z &= 0 &&\text{on } \partial\Omega. \end{aligned} \tag{2.4}$$

We notice that as long as $\nabla u^\delta_{\text{sm}}$, the gradient of the smoothed version u^δ_{sm} of noisy data u^δ, belongs to $L^\infty(\Omega)$, and s belongs to $L^2(\Omega)$, one can always seek the solution z to (2.4) in $H^1_0(\Omega)$, which leads to the compactness of the perturbed operator A on $L^2(\Omega)$, and which results in the ill-posedness of (2.3). Therefore, in this paper we consider $\bar{A}$ and A as operators acting from $L^2(\Omega)$ to itself and rely on an estimates of the form

$$\|\bar{r} - r^\delta\| = \|u - u^\delta_{\text{sm}}\| \leqslant \delta \quad \text{and} \quad \|\bar{A} - A\| \leqslant \varepsilon.$$

In general, the quantity ε can be larger than the estimate δ, i.e.,

$$\varepsilon \gg \delta,$$

as the following argument shows: Note that by (2.2) and (2.4), the operator $\bar{A} - A$ maps $s := a - a_0$ to the solution z of the Dirichlet problem

$$\begin{aligned} -\nabla(a_0\nabla z) &= \nabla(s\nabla(u - u^\delta_{\text{sm}})) \quad &&\text{in } \Omega, \\ z &= 0 &&\text{on } \partial\Omega. \end{aligned}$$

Now, by applying the Lax–Milgram theorem, there exists $c > 0$ such that

$$\|(\bar{A} - A)s\|_{H_0^1(\Omega)} \leqslant c\|\nabla(s\nabla(u - u_{\rm sm}^\delta))\|_{H_0^{-1}(\Omega)} = c\|s\nabla(u - u_{\rm sm}^\delta)\|_{L^2(\Omega)}.$$

Thus, we obtain

$$\|(\bar{A} - A)s\|_{L^2(\Omega)} \leqslant c\|s\nabla(u - u_{\rm sm}^\delta)\|_{L^2(\Omega)} \leqslant c\|s\|_{L^2(\Omega)}\|\nabla(u - u_{\rm sm}^\delta)\|_{L^\infty(\Omega)}.$$

Consequently,

$$\|\bar{A} - A\|_{L^2(\Omega)\to L^2(\Omega)} \leqslant c\|\nabla(u - u_{\rm sm}^\delta)\|_{L^\infty(\Omega)}.$$

As in [3], with additional assumptions, the above estimate can be strengthened to

$$\|\bar{A} - A\|_{L^2(\Omega)\to L^2(\Omega)} \leqslant c\|\nabla(u - u_{\rm sm}^\delta)\|_{L^2(\Omega)}.$$

Now, due to the ill-posedness of the operation $v \mapsto \nabla v$, the quantity $\|\nabla(u - u_{\rm sm}^\delta)\|_{L^2(\Omega)}$ can be much larger than $\|u - u_{\rm sm}\|_{L^2(\Omega)}$. For example, in [3], $\varepsilon = O(\sqrt{\delta})$. For the detailed method to obtain $u_{\rm sm}^\delta$ we refer the reader to [3, 15].

2.2. Descriptions of the MAM Algorithm

Once data mollification is completed, the operator equation (2.3) is to be solved numerically. Since (2.3) is ill-posed, we will apply Tikhonov regularization combined with the projection method. The Tikhonov regularized solution of (2.3) is obtained by solving the equation

$$\alpha s + A^*As = A^*r^\delta. \tag{2.5}$$

Since the MAM algorithm and the corresponding error estimation are applicable in a general setting, we assume, in this section, that $A : X \to Y$ is a compact operator between Hilbert spaces X and Y. For the linearized parameter identification problem as described in Section 2.1, we have $X = Y = L^2(\Omega)$. Now let $\{X_n : n \in \mathbb{N}\}$ be a sequence of finite-dimensional subspaces of X satisfying

$$\overline{\bigcup_{n\in\mathbb{N}} X_n} = X,$$

and for each $n \in \mathbb{N}$, let $P_n : X \to X$ be the orthogonal projection with range X_n. Then the projected version of (2.5) can be written as

$$\alpha s + P_nA^*AP_ns = P_nA^*r^\delta, \tag{2.6}$$

with its solution denoted by $s_{\alpha,n}^\delta$.

We assume that $X_n \subset X_{n+1}, n \in \mathbb{N}$. Let W_n be the orthogonal complement of X_{n-1} in X_n, $n \in \mathbb{N}$, that is, $X_n = X_{n-1} \oplus^\perp W_n$, where $W_0 := X_0$ and the notation $\oplus^\perp$ stands for the direct sum operation of two orthogonal subspaces. Thus, for any $k \in \mathbb{N}$ and $\ell \in \mathbb{N}$, we have the decomposition

$$X_{k+\ell} = X_k \oplus^\perp W_{k+1} \oplus^\perp W_{k+2} \oplus^\perp \cdots \oplus^\perp W_{k+\ell}. \tag{2.7}$$

Under this settings, the solution $s_{\alpha,k+\ell}^\delta \in X_{k+\ell}$ to (2.6) can be represented as

$$s_{\alpha,k+\ell}^\delta = (s_{\alpha,k+\ell}^\delta)_0 + \sum_{j=1}^{\ell}(s_{\alpha,k+\ell}^\delta)_j,$$

where $(s_{\alpha,k+\ell}^\delta)_0 \in X_k$ and $(s_{\alpha,k+\ell}^\delta)_j \in W_{k+j}$ for $j = 1, \ldots, \ell$.

As in [7], we use $w(i)$ to denote the dimension of the space W_i and assume that W_i has a basis $\{w_{ij} : j \in \mathbb{Z}_{w(i)}\}$, where $\mathbb{Z}_n := \{0, 1, \ldots, n-1\}$. Then according to (2.7),

$$X_n = \operatorname{span}\{w_{ij} : j \in \mathbb{Z}_{w(i)},\ i \in \mathbb{Z}_{n+1}\}.$$

Let $h_i := \max\{\operatorname{diam}(S_{ij}) : j \in \mathbb{Z}_{w(i)}\}$, where S_{ij} is the support set of w_{ij} and $\operatorname{diam}(S)$ denotes the diameter of a set $S \subset \mathbb{R}^d$. Denoting the dimension of the space X_n by d_n, we assume the following multiscale properties:

$$d_n \sim \mu^n, \quad w(i) \sim \mu^i, \quad h_i \sim \mu^{-i/d}, \tag{2.8}$$

where $\mu > 1$ is an integer. In the one-dimensional case, μ can be taken as 2, which can be viewed as the scaling factor in the construction of the wavelet basis. For details on the construction of such multiscale spaces we refer the reader to [9].

Now, in order to apply the multilevel augmentation to equation (2.6), we split the initial approximate solution $s^\delta_{\alpha,n}$ into two terms as

$$s^\delta_{\alpha,n} = s^{\delta,L}_{\alpha,n} + s^{\delta,H}_{\alpha,n}$$

with $s^{\delta,L}_{\alpha,n} \in X_k$ and $s^{\delta,H}_{\alpha,n} \in W_{k,\ell} := W_{k+1} \oplus^\perp W_{k+2} \oplus^\perp \cdots \oplus^\perp W_{k+\ell}$, which correspond to lower and higher resolutions of the approximate solution $s^\delta_{\alpha,n}$, respectively.

Defining $Q_{n+1} := P_{n+1} - P_n$, $n \in \mathbb{N}$, the MAM algorithm can be described as follows:

Algorithm for MAM.

1. *Fixing initial level:* For a fixed $k > 0$, solve (2.6) for $s^\delta_{\alpha,k} \in X_k$ with $n = k$, i.e.,

$$\alpha s^\delta_{\alpha,k} + P_n A^* A P_n s^\delta_{\alpha,k} = P_n A^* r^\delta,$$

and set $s^\delta_{\alpha,k,0} = s^\delta_{\alpha,k}$. For $\ell = 1, 2, \ldots$, suppose $s^\delta_{\alpha,k,\ell-1} \in X_{k+\ell-1}$ has been obtained, then $s^\delta_{\alpha,k,\ell}$ will be constructed by higher and lower resolution respectively, i.e.,

$$s^\delta_{\alpha,k,\ell} := s^{\delta,L}_{\alpha,k,\ell} + s^{\delta,H}_{\alpha,k,\ell}.$$

2. *Higher resolution part:* Compute

$$s^{\delta,H}_{\alpha,k,\ell} = \alpha^{-1}(P_{k+\ell} - P_k)(A^* r^\delta - A^* A s^\delta_{\alpha,k,\ell-1})$$

i.e.,

$$\underbrace{(s^\delta_{\alpha,k,\ell})_j}_{\in W_{k+j}} = \alpha^{-1} Q_{k+j}(A^* r^\delta - A^* A s^\delta_{\alpha,k,\ell-1}) \quad \text{for } j = 1, 2, \ldots, \ell.$$

3. *Lower resolution part:* Solve

$$P_k(\alpha I + A^* A) s^{\delta,L}_{\alpha,k,\ell} = P_k A^* r^\delta - P_k A^* A s^{\delta,H}_{\alpha,k,\ell}$$

for $s^{\delta,L}_{\alpha,k,\ell} \in X_k$, i.e.,

$$P_k(\alpha I + A^* A) \underbrace{(s^\delta_{\alpha,k,\ell})_0}_{\in X_k} = P_k A^* r^\delta - P_k A^* A \underbrace{\left(\sum_{j=1}^{l} (s^\delta_{\alpha,k,\ell})_j \right)}_{s^{\delta,H}_{\alpha,k,\ell}}.$$

As in [12], the MAM algorithm above can be rewritten in a matrix form. Then the iterative scheme described above can be realized by augmenting the corresponding matrices and vectors in each iteration and due to this, the algorithm takes its name. At the same time, if we introduce the notations

$$\mathcal{B}_{k,\ell}(\alpha) := I + \alpha^{-1} P_k A^* A P_{k+\ell} \quad \text{and} \quad \mathcal{C}_{k,\ell}(\alpha) := \alpha^{-1}(P_{k+\ell} - P_k) A^* A P_{k+\ell},$$

and formally, augment $s^\delta_{\alpha,k,\ell-1}$ into the form

$$\tilde{s}^\delta_{\alpha,k,\ell} := \begin{bmatrix} s^\delta_{\alpha,k,\ell-1} \\ 0 \end{bmatrix} \in X_{k+\ell},$$

then the solution $s^\delta_{\alpha,k,\ell}$ for the next level is obtained by solving the following equation with the augmented matrices and the corresponding vectors:

$$\mathcal{B}_{k,\ell}(\alpha) s^\delta_{\alpha,k,\ell} = \alpha^{-1} P_{k+\ell} A^* f^\delta - \mathcal{C}_{k,\ell}(\alpha) \tilde{s}^\delta_{\alpha,k,\ell}. \tag{2.9}$$

To estimate the accuracy of the augmented solution, we need the following assumptions for the considered inverse problem.

Assumption 2.1. *The solution $s = a - a_0$ to (2.1) belongs to the source set*

$$\mathcal{M}_{\phi,R} := \{ s \in X : s = \phi(\bar{A}^* \bar{A}) w,\ \|w\| \leqslant R \}, \tag{2.10}$$

where ϕ is an 'index function' defined on an interval $[0,b]$ containing the spectrum of $\bar{A}^*\bar{A}$, which in general is operator monotone (see [16, 17]) and satisfies the conditions $\phi(0) = 0$ and

$$\sup_{0<\lambda\leqslant b} \left| \frac{\alpha}{\alpha+\lambda} \right| \phi(\lambda) \leqslant c\phi(\alpha) \quad \textit{for all } \alpha \in (0, \bar{\alpha}]$$

for some $\bar{\alpha} > 0$.

It is worth pointing out that the well-known index functions related to Tikhonov regularization of ill-posed operator equations are of the forms

$$\begin{aligned} \phi(\lambda) &= \lambda^\nu, & \lambda > 0 \text{ for } 0 < \nu \leqslant 1, \\ \phi(\lambda) &= \log^{-p}(1/\lambda), & 0 < \lambda < 1 \text{ for } p > 0, \end{aligned}$$

which are contained in our considerations about index functions.

Assumption 2.2. *For some positive constant $r > 0$, there exist $c_1 > 0$ and $\mu > 1$ such that*

$$\|\bar{A}(I - P_n)\| \leqslant c_1 \mu^{-rn/d} \quad \textit{and} \quad \|A(I - P_n)\| \leqslant c_1 \mu^{-rn/d} \tag{2.11}$$

for all $n \in \mathbb{N}$.

The above assumption is motivated by the following considerations.

If $\bar{A}$ is defined as in Section 2.1, then $\bar{A}^*$ is given by

$$\bar{A}^* \psi = \nabla u \cdot \nabla \tilde{\psi}, \tag{2.12}$$

where $\tilde{\psi}$ solves the adjoint problem

$$\begin{aligned} -\nabla(a_0 \nabla \tilde{\psi}) &= \psi \quad \text{in } \Omega, \\ \tilde{\psi} &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

When $\bar{A}$ acts from $L^2(\Omega)$ to $L^2(\Omega)$, for every $\psi \in L^2(\Omega)$, $\bar{A}^*\psi$ will be an element in $H^1(\Omega)$ as long as ∇u belongs to $H^1(\Omega)$. In this case, when P_n is the orthogonal projection from $L^2(\Omega)$ onto the n-dimensional space X_n of piecewise linear continuous functions with multiscale properties (2.8), that is, the mesh size X_n is $\mu^{-n/d}$, then we have

$$\|I - P_n\|_{H_0^1(\Omega)\to L^2(\Omega)} \leqslant \tilde{c}\mu^{-rn/d}$$

according to the Jackson type inequality in approximation theory (cf. [1]). Then

$$\|\bar{A}(I - P_n)\| = \|(I - P_n)\bar{A}^*\| \leqslant \|I - P_n\|_{H_0^r(\Omega)\to L^2(\Omega)} \|\bar{A}^*\|_{L^2(\Omega)\to H^r(\Omega)} \leqslant c_r\mu^{-rn/d}$$

holds true with $r = 1$.

As for the second inequality in (2.11) concerning noisy operator A, one can use a similar argument by replacing ∇u in (2.12) by $\nabla u^\delta_{\mathrm{sm}}$. Then the inequality $\|A(I - P_n)\| \leqslant c_1\mu^{-rn/d}$ holds for some r in $(0, 1)$, when the data mollification can guarantee $\nabla u^\delta_{\mathrm{sm}} \in H^r(\Omega)$.

On the other hand, while considering singular value decomposition, in particular, if the same orthogonal system is used for the SVD of $\bar{A}$ and A, and the noise of operator A only embodies in the coefficients in SVD corresponding to different frequencies, then the two equalities in (2.11) will hold true for the same $r > 0$. An example in [3, Section 1] can illustrate this when $\bar{A}$ and A are defined by (2.2) and (2.4) on an interval $[0, 2\pi]$. We refer the reader to [6] for a general discussion, where noisy operators having similar properties are considered.

3. Estimation of Accuracy

Let

$$\bar{s}_\alpha = \bar{R}_\alpha \bar{r} \quad \text{and} \quad s^\delta_{\alpha,n} = R^{(n)}_\alpha r^\delta,$$

where

$$\bar{R}_\alpha = (\alpha I + \bar{A}^*\bar{A})^{-1}\bar{A}^* \quad \text{and} \quad R^{(n)}_\alpha = (\alpha I + P_nA^*AP_n)^{-1}P_nA^*.$$

Then we obtain

$$\|\bar{s}_\alpha - s^\delta_{\alpha,n}\| \leqslant \|(\bar{R}_\alpha - R^{(n)}_\alpha)\bar{r}\| + \|R^{(n)}_\alpha(\bar{r} - r^\delta)\| \leqslant \|(\bar{R}_\alpha - R^{(n)}_\alpha)\bar{r}\| + \frac{\varepsilon}{\sqrt{\alpha}}.$$

Hence, the proof of the following lemma is a consequence of the assumption (2.11) and [18, Corollary 5.1].

Lemma 3.1. *Under Assumption 2.2, the estimate*

$$\|\bar{s}_\alpha - s^\delta_{\alpha,n}\| \leqslant \frac{c_2}{\sqrt{\alpha}}(\varepsilon + \mu^{-rn/d}) \tag{3.1}$$

holds.

We denote the estimate in (3.1) by $\gamma_{\alpha,n}$, i.e.,

$$\gamma_{\alpha,n} := \frac{c_2}{\sqrt{\alpha}}(\varepsilon + \mu^{-rn/d}),$$

and note that

$$\frac{\gamma_{\alpha,n}}{\gamma_{\alpha,n+1}} \leqslant \sigma := \mu^{r/d} \quad \text{for } n \in \mathbb{N}.$$

The proof of the following lemma is similar to [12, Proposition 3.1]. In order to keep the present paper self-contained, we give a brief sketch of the proof.

Lemma 3.2. *Under Assumption 2.2, there exists an integer N such that for any $k \in \mathbb{N}$ with $k \geqslant N$ and for any $\ell \in \mathbb{N}$, the estimate*

$$\|s^\delta_{\alpha,k,\ell} - s^\delta_{\alpha,k+\ell}\| \leqslant \gamma_{\alpha,k+\ell} \tag{3.2}$$

holds, where $s^\delta_{\alpha,k,\ell}$ is obtained from the MAM algorithm and $s^\delta_{\alpha,k+\ell}$ solves (2.6) with $n = k+\ell$.

Proof. From (2.6) and (2.9), we have

$$\mathcal{B}_{k,\ell}(\alpha)(s^\delta_{\alpha,k,\ell} - s^\delta_{\alpha,k+\ell}) = \mathcal{C}_{k,\ell}(\alpha)(\tilde{s}^\delta_{\alpha,k,\ell} - s^\delta_{\alpha,k,\ell}). \tag{3.3}$$

From the definitions of operators $\mathcal{B}_{k,\ell}(\alpha)$ and $\mathcal{C}_{k,\ell}(\alpha)$, we obtain

$$\|\mathcal{C}_{k,\ell}(\alpha)\| = \|\alpha^{-1}(P_{k+\ell} - P_k)A^*AP_{k+\ell}\| \leqslant c\alpha^{-1}\mu^{-rk/d}$$

for some constant $c > 0$ so that

$$\|\mathcal{C}_{k,\ell}(\alpha)\| \to 0, \quad \text{uniformly for } \ell \in \mathbb{N} \text{ as } k \to \infty. \tag{3.4}$$

Also, for any $x \in X$

$$\|\mathcal{B}_{k,\ell}(\alpha)(x)\| = \big\|\big(I + \alpha^{-1}P_{k+\ell}A^*AP_{k+\ell} - \mathcal{C}_{k,\ell}(\alpha)\big)x\big\| \geqslant (1 - \|\mathcal{C}_{k,\ell}(\alpha)\|)\|x\|.$$

Therefore, there exists an integer N (depending on α) such that

$$\|\mathcal{B}^{-1}_{k,\ell}(\alpha)\| \leqslant \frac{1}{1 - \|\mathcal{C}_{k,\ell}(\alpha)\|} \tag{3.5}$$

for $k \geqslant N$. Now we can prove (3.2) by induction on index ℓ. When $\ell = 0$, since $s^\delta_{\alpha,k,0} = s^\delta_{\alpha,k}$, estimate (3.2) holds obviously. Suppose that (3.2) holds for $\ell = m-1$, $m \in \mathbb{N}^*$. Recalling the definition of $\tilde{s}^\delta_{\alpha,k,\ell}$, and from (3.1) and the induction hypothesis, we have

$$\begin{aligned}\|s^\delta_{\alpha,k+m} - \tilde{s}^\delta_{\alpha,k,m}\| &\leqslant \|s^\delta_{\alpha,k+m} - \bar{s}_\alpha\| + \|\bar{s}_\alpha - s^\delta_{\alpha,k+m-1}\| + \|s^\delta_{\alpha,k+m-1} - \tilde{s}^\delta_{\alpha,k+m}\| \\ &\leqslant \gamma^\delta_{\alpha,k+m} + 2\gamma^\delta_{\alpha,k+m-1} \\ &\leqslant (1+2\sigma)\gamma^\delta_{\alpha,k+m}.\end{aligned}$$

Then in view of (3.3) and (3.5), we can obtain

$$\|s^\delta_{\alpha,k,m} - s^\delta_{\alpha,k+m}\| \leqslant \|\mathcal{B}^{-1}_{k,m}(\alpha)\mathcal{C}_{k,m}(\alpha)\| \leqslant \frac{\|\mathcal{C}_{k,m}(\alpha)\|}{1 - \|\mathcal{C}_{k,m}(\alpha)\|}(1+2\sigma)\gamma^\delta_{\alpha,k+m}.$$

From (3.4) we notice that for fixed $\alpha > 0$, when k is sufficiently large,

$$\frac{\|\mathcal{C}_{k,m}(\alpha)\|}{1 - \|\mathcal{C}_{k,m}(\alpha)\|} \leqslant \frac{1}{1+2\sigma},$$

which implies that

$$\|s^\delta_{\alpha,k,m} - s^\delta_{\alpha,k+m}\| \leqslant \gamma^\delta_{\alpha,k+m}.$$

Thus, we have proved (3.2) for $\ell = m$. This completes the proof. □

Since in Assumption 2.1 we assume that our exact solution s belongs to the source set $\mathcal{M}_{\phi,R}$, we obtain the estimate corresponding to the Tikhonov regularized solution $\bar{s}_\alpha$ as

$$\|s - \bar{s}_\alpha\| \leqslant c_3\phi(\alpha) \tag{3.6}$$

(cf. [18, Theorem 4.3.1]). Based on (3.6), together with (3.1) and (3.2), we can get the final estimate for the regularized solution obtained by the MAM algorithm as follows.

Theorem 3.3. *Under the Assumptions 2.1 and 2.2, there exists an integer $N \in \mathbb{N}$ such that for any $k \in \mathbb{N}$ with $k \geqslant N$ and for any $\ell \in \mathbb{N}$*

$$\|s - s^\delta_{\alpha,k,\ell}\| \leqslant c_3\phi(\alpha) + 2c_2\Big(\frac{\varepsilon}{\sqrt{\alpha}} + \frac{\mu^{-r(k+\ell)/d}}{\sqrt{\alpha}}\Big). \tag{3.7}$$

4. Parameter Choice by Balancing Principle

We first choose initial level $k := k(\alpha) > N$ large enough such that for any $l \in \mathbb{N}$

$$\mu^{-r(k+\ell)/d} \leqslant \varepsilon. \tag{4.1}$$

Then, in view of (3.7), we have

$$\|s - s^\delta_{\alpha,k(\alpha),\ell}\| \leqslant c_3\phi(\alpha) + K\frac{\varepsilon}{\sqrt{\alpha}}, \tag{4.2}$$

where K, for example, can be taken as $4c_2$. Observing that the two terms in estimate (4.2) have different monotonic properties with respect to α, and the function ϕ in general is unknown, we employ an posteriori parameter choice rule based on a balancing principle. In this adaptive strategy, we follow an extended form of the balancing principle developed in [4,5]. This modified principle adapts to a suitable estimate for constant K in (4.2).

To describe such a strategy, we introduce the set

$$\Delta_I := \big\{\alpha_i = a_0 q^i : i = 0, 1, \ldots, I\big\},$$

with $\alpha_0 = \varepsilon^2$, $q > 1$ and I such that $\alpha_{I-1} \leqslant 1 \leqslant \alpha_I$. The regularization parameter will be chosen from the finite set above.

For any given value of the constant K in (4.2), one can select $\alpha = \alpha(K)$ by the following adaptive rule:

$$\alpha(K) = \max\Big\{\alpha_j \in \Delta_I : \|s^\delta_{\alpha_i,k(\alpha_i),\ell} - s^\delta_{\alpha_j,k(\alpha_j),\ell}\| \leqslant K\varepsilon\Big(\frac{3}{\sqrt{\alpha_j}} + \frac{1}{\sqrt{\alpha_i}}\Big),\ i = 0, 1 \ldots, j-1\Big\}. \tag{4.3}$$

Since the value of K cannot be precisely determined, we combine the adaptive strategy (4.3) with successive testing of the hypothesis that K is not larger than some term in a fixed geometric sequence

$$\mathcal{K}_M := \big\{K_m = K_0 p^m : m = 0, 1, \ldots, M\big\}, \tag{4.4}$$

with $p > 1$. For each of these hypotheses, $\alpha = \alpha(K_m) \in \Delta_I$ is chosen according to the rule (4.3), which results in a nondecreasing sequence,

$$\alpha(K_0) \leqslant a(K_1) \leqslant \cdots \leqslant a(K_m) \leqslant \cdots \leqslant a(K_M).$$

We further assume a two-sided stability bound

$$\tilde{p}K\frac{\varepsilon}{\sqrt{\alpha}} \leqslant \|\bar{s}_\alpha - s^\delta_{\alpha,k(\alpha),\ell}\| \leqslant K\frac{\varepsilon}{\sqrt{\alpha}} \tag{4.5}$$

for some $\tilde{p} \in (0,1)$ and for any $\alpha \in \Delta_I$, where $\bar{s}_\alpha$ is defined in Lemma 3.1. At the same time, we require that the testing set (4.4) is designed in such a way that there are two adjacent terms $K_{\bar{m}}, K_{\bar{m}+1} \in \mathcal{K}_M$ such that

$$K_{\bar{m}} \leqslant \tilde{p}K < K \leqslant K_{\bar{m}+1}, \tag{4.6}$$

which means that the term $K_{\bar{m}+1}$ with an unknown index $\bar{m}+1$ is the best candidate for the estimate to K among the elements in $\mathcal{K}_M$. Then, as in [5], one can show when K_m is strictly smaller than the unknown lower bound $K_{\bar{m}}$ in (4.6), i.e., $m \leqslant \bar{m}-1$, then $\alpha(K_m)$ is smaller than a threshold determined by α_0 and p. Therefore, we have the following proposition.

Proposition 4.1. *Let assumptions* (4.5) *and* (4.6) *hold, and let*

$$\alpha(K_{\tilde{m}}) := \min\Big\{\alpha(K_m) : \alpha(K_m) \leqslant 9\alpha_0\Big(\frac{p^2+1}{p-1}\Big)^2,\ m = 0, 1, \ldots, M\Big\}.$$

Then either $\tilde{m} = \bar{m}$ *or* $\tilde{m} = \bar{m}+1$.

In order to guarantee the regularized solution stable enough, the final choice of parameter α given by the balancing principle is

$$\alpha_+ = \alpha(K_{\tilde{m}+1}).$$

Theorem 4.2. *Under assumptions* (4.5) *and* (4.6), *the estimate*

$$\|s - s^\delta_{\alpha_+,k(\alpha_+),l}\| \leqslant 6p^2\sqrt{q}c_3\phi(\theta^{-1}(K\varepsilon)) \tag{4.7}$$

holds, where $\theta(\lambda) := c_3\phi(\lambda)\sqrt{\lambda}$.

Proposition 4.1 and Theorem 4.2 can be proved by taking the additional parameter κ as 1 in the counterparts in [5] or fixing the stability exponent ν as 1/2 in [4].

Theorem 4.2 suggests that α_+ renders a regularized solution $s^\delta_{\alpha_+,k(\alpha_+),\ell}$ with order optimal accuracy. At the same time, we obtain $K_{\tilde{m}+1}$ as a reliable estimate to constant K. If the index function ϕ in source condition (2.10) is taken as $\phi(\lambda) = c\lambda^\nu$, $0 < \nu \leqslant 1$, then one can prove that, under the parameter chosen by the balancing principle, the estimate (4.7) coincides with the classical rate for Tikhonov regularization,

$$\|s - s^\delta_{\alpha_+,k(\alpha_+),l}\| \leqslant c\varepsilon^{\frac{2\nu}{2\nu+1}}.$$

5. Numerical Tests

In this section, we present numerical test results for two examples given in [14] to illustrate the results of the above sections. Here we use MATLAB-code in the one-dimensional case, where $\Omega = (0,1)$, and the equations in (1.1) remain as

$$(au_x)_x = 0 \text{ in } (0,1), \quad u(0) = 0, \quad u(1) = 1.$$

We fix initial guess $a_0 \equiv 1$, which implies $u_0(x) = x$.

Example 5.1.

$$a(x) = \begin{cases} 1 + \frac{1}{3}\sin^2(\pi\frac{x-0.5}{0.2}) & \text{if } x \in [0.3, 0.7], \\ 1 & \text{otherwise.} \end{cases}$$

$$u(x) = \begin{cases} \frac{x}{1-0.2(2-\sqrt{3})} & \text{if } x \in [0, 0.3], \\ \frac{0.3+\frac{0.2\sqrt{3}}{2\pi}\left(\arctan(\sqrt{3}\tan(\frac{\pi}{2}\frac{x-0.5}{0.2}))+\arctan(\frac{1}{\sqrt{3}}\tan(\frac{\pi}{2}\frac{x-0.5}{0.2}))+\pi\right)}{1-0.2(2-\sqrt{3})} & \text{if } x \in [0.3, 0.7], \\ \frac{x-0.2(2-\sqrt{3})}{1-0.2(2-\sqrt{3})} & \text{if } x \in [0.7, 1], \end{cases}$$

Example 5.2.

$$a(x) = \begin{cases} 1 + \sin(\frac{\pi}{2}\frac{x-0.5}{0.45} + 1) & \text{if } x \in [0.05, 0.95], \\ 1 & \text{otherwise.} \end{cases}$$

$$u(x) = \begin{cases} \frac{x}{1-0.45(2-\frac{4}{\pi})} & \text{if } x \in [0, 0.05], \\ \frac{0.05+\frac{0.9}{\pi}(\tan(\frac{\pi}{4}\frac{x-0.5}{0.45}))}{1-0.45(2-\frac{4}{\pi})} & \text{if } x \in [0.05, 0.95], \\ \frac{x-0.45(2-\frac{4}{\pi})}{1-0.45(2-\frac{4}{\pi})} & \text{if } x \in [0.95, 1], \end{cases}$$

In fact, for this special case, since $a_0 \equiv 1$, the implicit definition of operator A in (2.4) can be rewritten in an explicit way as

$$(As)(x) = -\int_0^x s(t)(u^\delta_{\text{sm}}(t))'dt + x\int_0^1 s(t)(u^\delta_{\text{sm}}(t))'dt$$

so that the considered numerical tests have a similar feature to the Volterra integral equation of the second type. For this reason, we take the same wavelets basis functions as in [8,11,12]. In this setting, we choose basis functions for the X_0 as

$$w_{00}(t) = 2 - 3t \quad \text{and} \quad w_{01}(t) = -1 + 3t,$$

and basis functions for the space W_1 as

$$w_{10}(t) = \begin{cases} 1 - \frac{9}{2}t & \text{if } t \in [0, \frac{1}{2}], \\ -1 + \frac{3}{2}t & \text{if } t \in [\frac{1}{2}, 1], \end{cases} \quad \text{and} \quad w_{11}(t) = \begin{cases} \frac{1}{2} - \frac{3}{2}t & \text{if } t \in [0, \frac{1}{2}], \\ -\frac{7}{2} + \frac{9}{2}t & \text{if } t \in [\frac{1}{2}, 1]. \end{cases}$$

Then W_i, $i = 2, 3, \ldots, n$ can be recursively constructed. In the tests, we put additional noise directly to the right-hand side r of the linearized equation, i.e., we construct $r^\delta = r + \delta\xi$, where ξ has uniformly distributed random values with $\|\xi\| \leqslant 1$ and $\delta = \|r\|\tilde{\delta}$ with $\tilde{\delta} = 0.05$ or 0.1. The data mollification is done by piecewise linear interpolation. As in [3], we have the noise level $\varepsilon \sim \sqrt{\delta}$. The test results are summarized in Tables 1 and 2.

From the results obtained in Section 3 and 4, one can find that the suitable initial level k for the MAM algorithm depends on α. When α increases, the corresponding regularized problem becomes more stable and correspondingly k can be chosen relatively smaller. However, by far there are no theoretical results suggesting specific (quantized) relations between k and α. In the numerical test, we fix the initial level as $k = 16$ and $k = 32$. In these cases, one can observe that the simulated solutions $s^\delta_{\alpha,k,\ell}$ obtained by MAM methods have

$\tilde{\delta}$	k	ℓ	$\|s - s^{\delta}_{\alpha,k,\ell}\|$	$\|s - s^{\delta}_{\alpha,k+\ell}\|$	α_+	K
5%	16	48	0.0560	0.0561	0.017	17.44
		112	0.0495	0.0492	0.0012	2.14
	32	32	0.0262	0.0260	3×10^{-4}	23.23
		96	**0.0212**	0.0190	2×10^{-4}	0.67
10%	16	48	0.0613	0.0612	0.0023	10.83
		112	0.0627	0.0625	0.0025	1.21
	32	32	0.0480	0.0479	0.0011	0.47
		96	0.0478	0.0477	0.011	0.51

Table 1. Test results for Example 5.1.

$\tilde{\delta}$	k	ℓ	$\|s - s^{\delta}_{\alpha,k,\ell}\|$	$\|s - s^{\delta}_{\alpha,k+\ell}\|$	α_+	K
5%	16	48	0.0036	0.0034	0.0079	1.74
		112	0.0030	0.0028	0.0072	2.55
	32	32	0.0022	0.0022	0.0034	6.80
		96	**0.0020**	0.0020	0.0042	2.38
10%	16	48	0.0040	0.0038	0.0096	0.90
		112	0.0046	0.0044	0.0106	1.31
	32	32	0.0039	0.0037	0.0097	1.85
		96	0.0088	0.0088	0.0194	0.51

Table 2. Test results for Example 5.2.

very slight differences with $s^{\delta}_{\alpha,k+\ell}$, which verifies the estimation in Lemma 3.2. On the other hand, in the calculations of MAM, one only needs to invert a stiff matrix with size determined by initial level $k = 16$ or $k = 32$, instead of a large matrix with size $k + \ell = 64$ or $k + \ell = 128$ obtained by direct discretization. The remaining computations in the MAM algorithm are the multiplications of matrices and vectors, which are well-posed. In this sense, the computational cost is dramatically reduced. Besides the numerical illustration presented in this paper, the MAM algorithm will become more instructive when higher dimensional numerical applications are considered.

During the process of the parameter choice balancing principle, we can also obtain an estimate $K_{\tilde{m}+1}$ of constant K in (4.2) in an adaptive way. From (4.1) and (3.7), one can find that K combines the influence of both initial noise level and discretization level. Therefore, it varies with quantities δ ($\tilde{\delta}$), k and ℓ. The variation of K has also been detected by the balancing principle as shown in Tables 1 and 2.

Figure 1 shows one of the best simulations to coefficient $a(x)$ in Example 5.1 with $k = 32$, $\ell = 96$ and $\alpha_+ = 2 \times 10^{-4}$.

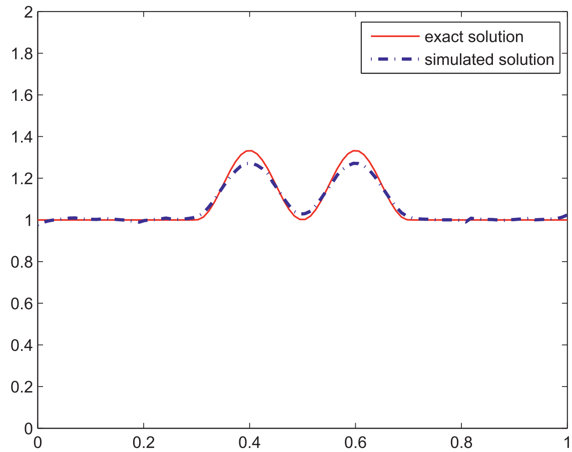

Figure 1. A simulated solution $s^{\delta}_{\alpha,k,\ell} + a_0$ in Example 5.1 with $\tilde{\delta} = 5\%$, $k = 32$, $\ell = 96$ and $\alpha_+ = 2 \times 10^{-4}$.

Acknowledgments

This research was partially supported by the Natural Science Foundation of China (no. 11201497), the Research Fund for the Doctoral Program of Higher Education of China (20120171120005), Guangdong Provincial Government of China through the "Computational Science Innovative Research Team" program and Guangdong Province Key Laboratory of Computational Science at the Sun Yat-sen University.

References

[1] R. A. Adams and J. F. John, *Sobolev Spaces*, 2nd edn., Pure Appl. Math. (Amst.), 140, Elsevier, Amsterdam, 2003.

[2] C. Bardos, D. Brézis, and H. Brezis, *Perturbations singulières et prolongements maximaux d'opérateurs positifs*, Arch. Rational Mech. Anal., **53** (1973/74), pp. 69–100.

[3] H. Cao, *Discretized Tikhonov-Phillips regularization for a naturally linearized parameter identification problem*, J. Complexity, **21** (2005), pp. 864–877.

[4] H. Cao, M. C. Klibanov, and S. V. Pereverzev, *A Carleman estimate and the balancing principle in the quasi-reversibility method for solving the Cauchy problem for the Laplace equation*, Inverse Problems, **25** (2009), article ID 035005.

[5] H. Cao and S. Pereverzev, *Natural linearization for the identification of a diffusion coefficient in a quasi-linear parabolic system from short-time observations*, Inverse Problems, **22** (2006), pp. 2311–2330.

[6] L. Cavalier and N. W. Hengartner, *Adaptive estimation for inverse problems with noisy operators*, Inverse Problems, **21** (2005), pp. 1345–1361.

[7] Z. Chen, S. Ding, and H. Yang, *Multilevel augmentation algorithms based on fast collocation methods for solving ill-posed integral equations*, Comput. Math. Appl., **62** (2011), pp. 2071–2082.

[8] Z. Chen, C. A. Micchelli, and Y. Xu, *A multilevel method for solving operator equations*, J. Math. Anal. Appl., **262** (2001), pp. 688–699.

[9] Z. Chen, B. Wu, and Y. Xu, *Multilevel augmentation methods for solving operator equations*, Numer. Math. J. Chin. Univ., **14** (2005), pp. 31–55.

[10] Z. Chen, B. Wu, and Y. Xu, *Multilevel augmentation methods for differential equations*, Adv. Comput. Math., **24** (2006), pp. 213–238.

[11] Z. Chen, B. Wu, and Y. Xu, *Fast numerical collocation solutions of integral equations*, Commun. Pure Appl. Anal., **6** (2007), pp. 643–666.

[12] Z. Chen, Y. Xu, and H. Yang, *A multilevel augmentation method for solving ill-posed operator equations*, Inverse Problems, **22** (2006), pp. 155–174.

[13] Z. Chen, Y. Xu, and H. Yang, *Fast collocation methods for solving ill-posed integral equations of the first kind*, Inverse Problems, **24** (2008), article ID 065007.

[14] B. Kaltenbacher, *A projection-regularized Newton method for nonlinear ill-posed problems and its application to parameter identification problems with finite element discretization*, SIAM J. Numer. Anal., **37** (2000), no. 6, pp. 1885–1908.

[15] B. Kaltenbacher and J. Schöberl, *A saddle point variational formulation for projection-regularized parameter identification*, Numer. Math., **91** (2002), pp. 675–697.

[16] P. Mathé and S. Pereverzev, *Discretization strategy for linear ill-posed problems in variable Hilbert scales*, Inverse Problems, **19** (2003), no. 6, pp. 1263–1277.

[17] P. Mathé and S. Pereverzev, *Geometry of linear ill-posed problems in variable Hilbert scales*, Inverse Problems, **19** (2003), no. 3, pp. 789–803.

[18] M. T. Nair, *Linear Operator Equations: Approximation and Regularization*, World Scientific, Singapore, 2009.

[19] M. T. Nair and S. Pereverzev, *Regularized collocation method for Fredholm integral equations of the first kind*, J. Complexity, **23** (2007), pp. 454–467.

[20] S. Pereverzev and E. Schock, *On the adaptive selection of the parameter in regularization of ill-posed problems*, SIAM J. Numer. Anal., **43** (2005), pp. 2060–2076.

[21] U. Tautenhahn, *A fast iterative method for solving regularized parameter identification problems in elliptic boundary value problem*, Computing, **43** (1989), no. 1, pp. 47–58.