Advances in Applied Mathematics and Mechanics
Adv. Appl. Math. Mech., Vol. **8**, No. 1, pp. 1-18

DOI: 10.4208/aamm.2014.m722
February 2016

An Inverse Source Problem with Sparsity Constraint for the Time-Fractional Diffusion Equation

Zhousheng Ruan[1,2], Zhijian Yang[1] and Xiliang Lu[1,*]

[1] *School of Mathematics and Statistics, Wuhan University, Wuhan 430072, China*
[2] *School of Science, East China Institute of Technology, Nanchang 330013, China*

Received 17 August 2014; Accepted (in revised version) 28 November 2014

Abstract. In this paper, an inverse source problem for the time-fractional diffusion equation is investigated. The observational data is on the final time and the source term is assumed to be temporally independent and with a sparse structure. Here the sparsity is understood with respect to the pixel basis, i.e., the source has a small support. By an elastic-net regularization method, this inverse source problem is formulated into an optimization problem and a semismooth Newton (SSN) algorithm is developed to solve it. A discretization strategy is applied in the numerical realization. Several one and two dimensional numerical examples illustrate the efficiency of the proposed method.

AMS subject classifications: 65N21, 49M15

Key words: Inverse source problem, time-fractional diffusion equation, sparse constraint, elastic-net regularization method, semismooth Newton method.

1 Introduction

Let $\Omega \subseteq \mathbb{R}^d$, $d = 1,2,3$ be an open bounded domain with C^1-boundary, we consider the following time-fractional diffusion equation with homogeneous Dirichlet boundary condition:

$$\begin{cases} \dfrac{\partial^\alpha u(x,t)}{\partial t^\alpha} = \Delta u + f(x), & (x,t) \in \Omega \times (0,T), \\ u(x,t) = 0, & (x,t) \in \partial\Omega \times (0,T], \\ u(x,0) = u_0(x), & x \in \Omega. \end{cases} \tag{1.1}$$

*Corresponding author.
Email: zhshruan@whu.edu.cn (Z. S. Ruan), zjyang.math@whu.edu.cn (Z. J. Yang), xllv.math@whu.edu.cn (X. L. Lu)

 ©2016 Global Science Press

https://doi.org/10.4208/aamm.2014.m722 Published online by Cambridge University Press

The fractional derivative $\partial^\alpha u(x,t)/\partial t^\alpha$ is the Caputo fractional derivative which is defined by

$$\frac{\partial^\alpha u(x,t)}{\partial t^\alpha} = \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-\eta)^{-\alpha} \frac{\partial u}{\partial \eta} \mathrm{d}\eta, \quad 0<\alpha<1, \tag{1.2}$$

where $\Gamma(1-\alpha)$ is the Gamma function. The time-fractional diffusion equation has been successfully applied in many fields. For instance, in [24] it is applied to describe the diffusion in fractional geometry. The time-fractional diffusion is also closely related to a non-Markovian diffusion process [22] or continuous time random walks on fractals [27]. A comprehensive review on it can be found in [3], see also [15, 25].

In last decades, the mathematical analysis and the numerical realization to the time-fractional diffusion equation have been studied in many literatures, see e.g., [4, 10, 12, 18, 19, 23, 28] and references cited there. Meanwhile, many scholars consider the inverse problem corresponding to time-fractional diffusion equation. For example, based on the eigenfunction expansion and the Gel'fand-Levitan theory, the uniqueness of identifying the order of the fractional derivative and diffusion coefficient was established in Chen et al. [1] for one-dimensional time fractional diffusion equation. In [14], Jin and Rundell proposed an algorithm of the quasi-Newton type to reconstruct a spatially varying potential term in a one-dimensional time-fractional diffusion equation and the unique identifiability of the inverse problem had been established in the case where the time is sufficiently large and the set of input sources forms a complete basis in $L^2(0,1)$. Liu and Yamamoto proposed a numerical scheme for the backward problem based on the quasi-reversibility method in [20] and derived error estimates for the approximation under a priori smoothness assumption on the initial condition. In [34], Wang and Liu applied the data regularizing technique to deal with the backward problem, under the a priori information about the bound on initial function, the Hölder convergence result was established. An regularization method was proposed to solve a time fractional order backward heat conduction problem and the optimal stability error estimation was obtained in Xiong et al. [37]. Ye and Xu proposed a time-space spectral approximation algorithm based on the optimal control framework to solve the backward problem and they obtained a priori error estimate for the spectral approximation in [38]. By applying the separation of variables, Wang et al. [33] reconstructed a space-dependent source for the time-fractional diffusion equation by Tikhonov regularization method and provided the convergence estimates under an a priori and a posteriori parameter choice rule. In [36], Wei and Zhang transformed the time-dependent inverse source problem into a first kind Volterra integral equation and used a boundary element method combined with a generalized Tikhonov regularization to solve the Volterra integral equation. Based on the method of the eigenfunction expansion, the uniqueness of the inverse problem was proved by analytic continuation and Laplace transform in Zhang and Xu [39]. Zheng and Wei considered Fourier regularization method to solve the sideway problem for the time fractional advection-dispersion equation in a quarter plane in [40] and they obtained the convergence under a priori bound assumptions for the exact solution.

https://doi.org/10.4208/aamm.2014.m722 Published online by Cambridge University Press

In these studies, the reconstructed solution is assume to be smooth and some algorithms work for one dimensional spatial domain but not easy to generalize to two (or three) dimensional spatial domain. In this work, we consider the reconstruction of the time independent source term $f(x)$ by the noisy final time measurement $g^\delta(x)\approx u(x,T;f)$, with a priori sparsity constraint for the source $f(x)$. The sparsity of $f(x)$ is understood in the sense that supp$f(x)$ is small, which is a reasonable assumption when we consider the (nearly) pointed source term. To handle the sparse constraint of $f(x)$, we propose an elastic-net type regularization formulation as:

$$\min \frac{1}{2}\|u(x,T;f)-g^\delta(x)\|^2+\beta\|f\|_{L^1(\Omega)}+\frac{\gamma}{2}\|f\|^2, \tag{1.3}$$

where $u(x,t;f)$ is the solution of Eq. (1.1) with source term f. The elastic-net regularization has been widely applied for inverse problem in sequence space [13, 41] and the optimal control in functional space [2, 29, 30] Due to the superposition principle to the Eq. (1.1), we may assume the initial data $u_0(x)=0$ in later discussion. Then for any given source function $f(x)$, by solving Eq. (1.1) we can formally define a forward linear operator

$$M: f(x)\mapsto u(x,T;f), \tag{1.4}$$

and the optimization problem (1.3) can be equivalently rewritten as

$$\min J_{\beta,\gamma}(f)=\frac{1}{2}\|Mf-g^\delta(x)\|^2+\beta\|f\|_{L^1(\Omega)}+\frac{\gamma}{2}\|f\|^2. \tag{1.5}$$

The regularization problem (1.5) is a nonsmooth optimization problem, we will apply a semismooth Newton method to solve it. For a large class of nonsmooth optimization problems, such as the optimal control problem with control/state constraints, the optimal control problem with L^1-cost functional or the inverse problem with sparse constraint, the semismooth Newton method was known to be an efficient, locally superlinearly convergent technique; see e.g., [8, 9, 16, 17, 30]. Recall the notation of Newton differentiability: Let X and Y be two Banach spaces, with D an open set in X and a map F from D to Y.

Definition 1.1. $F: D\mapsto Y$ is called Newton (slantly) differentiable at $x\in D$, if there exists an open neighborhood $N_x\subset D$ and mappings $G: N_x\mapsto \mathcal{L}(X,Y)$ (bounded linear operator from X to Y) such that

$$\lim_{\|h\|_X\to 0}\frac{\|F(x+h)-F(x)-G(x+h)(h)\|_Y}{\|h\|_X}=0. \tag{1.6}$$

The mapping G is called the Newton derivative of F at x.

For a Newton differentiable (in properly chosen function space) operator equation $F(x)=0$, the semismooth Newton method is Newton type algorithm which replaces the Frechét derivative by Newton derivative during each iteration. In the function space setting, we will prove the locally superlinear convergence of semismooth Newton method to problem (1.5).

https://doi.org/10.4208/aamm.2014.m722 Published online by Cambridge University Press

Problem (1.5) is an infinite-dimensional optimization problem. Both an optimization algorithm and a discretization of the governing equation are needed in the numerical realization. There exist two different approaches: "optimization-then-discretization" and "discretization-then-optimization". Since semismooth Newton method has been given in the function space setting, it is natural to apply "optimization-then-discretization" approach. However during each Newton iteration, one needs to solve a coupled fractional diffusion system, which involves all unknowns in both spatial and temporal direction. This leads to a huge linear system, which is very expensive to solve it for every Newton iteration. To overcome this difficulty, we introduce a different strategy for discretization. More precisely, the whole algorithm in this article is divided into two steps. Firstly we compute the explicit form of the discrete operator $M_{h,\tau}$, i.e., given a basis $\{\varphi_i\}$ of the discretization space $W_{\tilde{h}}$ of the source term f, then for any φ_i we compute $M_{h,\tau}\varphi_i$ by solving a discrete fractional diffusion equation, see also [35]. This step is quite expensive since a lot of discrete fractional diffusion equations need to be solved, but it can be computed in a parallel manner. Once the explicit expression of $M_{h,\tau}$ is derived, we can discrete the Newton iteration by using the explicit form $M_{h,\tau}$. In general the discrete operator $M_{h,\tau}$ is a full matrix. Thanks to the a priori sparse information, by using the standard P_1 finite element discretization, the discrete optimal solution is also sparse. Therefore the semismooth Newton method (or equivalently primal dual active set method) is very attractive: at each iteration one needs to solve a least square problem only on a small size subset (on the active set). Combining with a continuation strategy, the second step (semismooth Newton iteration) is very efficient, see also [7, 31]. It should be noticed that to determine the regularization parameter, one needs to solve the optimization problem (1.5) several times with different parameters. By using the explicit form of the discrete operator $M_{h,\tau}$ and discrete semismooth Newton algorithm, it is quite cheap to solve problem (1.5) with given parameters β and γ.

The rest of this paper is organized as follows. The notations are introduced at the end of this section. In Section 2, the elastic-net regularization is introduced, the semismooth Newton method is given and its local superlinear convergence is proved. The discretization of governing equation and a discretization algorithm is proposed in Section 3. Several numerical examples are given in Section 4 to show the efficiency of the proposed algorithm.

The standard Sobolev spaces $L^q(\Omega)$, $q \geq 1$ and $H^s(\Omega)$, $s \in \mathbb{N}$ are equipped with the norm

$$\|f\|_{L^q(\Omega)} = \Big(\int_\Omega |f(x)|^q dx \Big)^{1/q}, \quad \|f\|_{H^s(\Omega)} = \Big(\sum_{|\alpha| \leq s} \Big\| \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}} \Big\|^2_{L^2(\Omega)} \Big)^{1/2},$$

where Ω can be omitted without confusion and $\|\cdot\|_{L^2(\Omega)}$ is denoted by $\|\cdot\|$ for simplicity. The inner product in $L^2(\Omega)$ induced by its norm reads

$$\langle f, g \rangle = \int_\Omega f(x) g(x) dx.$$

https://doi.org/10.4208/aamm.2014.m722 Published online by Cambridge University Press

Throughout the paper, the constant C and c are generic constants that might be different at different places.

2 Elastic-net regularization and semismooth Newton method

2.1 Elastic-net regularization

For any given $f(x)\in L^2(\Omega)$, recall the operator $M: f(x)\mapsto u(x,T;f)$, where $u(x,t;f)$ is the unique solution of Eq. (1.1) with initial condition $u_0=0$. Let $\{\lambda_k\}_{k=1}^{\infty}$ $(0<\lambda_1\le\lambda_2\le\cdots, k=1,\cdots,\infty)$, $\{\chi_k\}_{k=1}^{\infty}$ be the eigenvalues and the L^2 orthonormal eigenfunctions of negative Laplace operator $-\Delta$ with homogeneous Dirichlet boundary condition, respectively. Then Mf can be expressed as (see [28]):

$$Mf(x)=u(x,T;f)=\sum_{k=1}^{\infty}\Big(\int_0^T(T-\tau)^{\alpha-1}E_{\alpha,\alpha}(-\lambda_k(T-\tau)^{\alpha})\mathrm{d}\tau\Big)\langle f,\chi_k\rangle\chi_k, \tag{2.1}$$

where $E_{\alpha,\theta}(z)$ is the double-parameter Mittag-Leffler function:

$$E_{\alpha,\theta}(z):=\sum_{k=0}^{\infty}\frac{z^k}{\Gamma(\alpha k+\theta)},\quad z\in\mathbb{C},\quad \alpha>0,\quad \theta\ge 0.$$

With the help of the above expression of M and the property of the Mittag-Leffler function (see e.g., [15]), we can prove the property of the forward operator M.

Lemma 2.1. *M is a self-adjoint injective bounded linear operator from $L^2(\Omega)$ to $L^2(\Omega)$. Moreover $c\|f\|\le\|Mf\|_{H^2}\le C\|f\|$ holds for some $0<c<C<+\infty$ and hence M is a compact operator.*

Proof. Due to (2.1) and the property of Mittag-Leffler function [15], the solution $u(x,T;f)$ can be represented by

$$\begin{aligned}Mf(x)=u(x,T;f)&=\sum_{k=1}^{\infty}\Big(\int_0^T(T-\tau)^{\alpha-1}E_{\alpha,\alpha}(-\lambda_k(T-\tau)^{\alpha})\mathrm{d}\tau\Big)\langle f,\chi_k\rangle\chi_k\\&=\sum_{k=1}^{\infty}\frac{1}{\lambda_k}(1-E_{\alpha,1}(-\lambda_k T^{\alpha}))\langle f,\chi_k\rangle\chi_k.\end{aligned}$$

By the orthogonality of $\{\chi_k\}_{k=1}^{\infty}$ and the complete monotonicity of the function $E_{\alpha,1}(-t)$ [26] and the fact that $0\le E_{\alpha,1}(-t)\le 1$ for $t\in[0,+\infty)$, we can deduce that the operator M is self-adjoint, injective and bounded linear operator. Moreover

$$\|\Delta u(\cdot,t)\|_{L^2}^2=\sum_{k=1}^{\infty}\lambda_k^2\Big(\frac{1-E_{\alpha,1}(-\lambda_k t^{\alpha})}{\lambda_k}\Big)^2\langle f,\chi_k\rangle^2=\sum_{k=1}^{\infty}(1-E_{\alpha,1}(-\lambda_k t^{\alpha}))^2\langle f,\chi_k\rangle^2,$$

https://doi.org/10.4208/aamm.2014.m722 Published online by Cambridge University Press

and

$$\|f\|^2=\sum_{k=1}^{\infty}\langle f,\chi_k\rangle^2\geq\sum_{k=1}^{\infty}(1-E_{\alpha,1}(-\lambda_k t^\alpha))^2\langle f,\chi_k\rangle^2$$
$$\geq\sum_{k=1}^{\infty}(1-E_{\alpha,1}(-\lambda_1 t^\alpha))^2\langle f,\chi_k\rangle^2=(1-E_{\alpha,1}(-\lambda_1 t^\alpha))^2\|f\|^2.$$

Combining above inequality with standard elliptic estimation, there exist positive constants c, C such that

$$c\|f\|\leq\|u(x,T;f)\|_{H^2}\leq C\|f\|. \tag{2.2}$$

Then compactness of the map M follows by Sobolev compact embedding theorem. □

From Lemma 2.1, we find the forward operator M is a smoothing operator and hence the associated inverse problem is ill-posed. By exploiting the sparse structure of the solution, we introduce the elastic-net regularization and obtain the nonsmooth optimization problem (1.5). The two regularization parameters β and γ are used to balance the fidelity term, sparsity level and smoothness of the solution. Since the operator M is linear, $\beta\|f\|_{L^1}$ is convex, the cost functional $J_{\beta,\gamma}(f)$ is strongly convex. Therefore the unique existence of the minimizer of (1.5) can be obtained by standard arguments, see e.g., [5].

Theorem 2.1. *For any positive β, γ, there exists a unique minimizer to problem* (1.5).

2.2 Semismooth Newton algorithm

We first derive the optimality system of problem (1.5). Let f^* be the optimal solution of problem (1.5), then there exists $d^*\in\partial\|\cdot\|_{L^1}|_{f^*}$ which satisfies

$$M^t(Mf^*-g^\delta)+\beta d^*+\gamma f^*=0.$$

One may observe that $d^*\in\partial\|\cdot\|_{L^1}|_{f^*}\Leftrightarrow f^*=T_c(f^*+cd^*)$ for any positive constant c (see [5]), where $T_c(\cdot)$ is the soft-thresholding operator with $T_c(f)=sgn(f)\max(|f|-c,0)$, see Fig. 1. By introducing the adjoint variable p^*, the optimality system can be written as

$$p^*=-M^t(Mf^*-g^\delta),\quad p^*=\beta d^*+\gamma f^*,\quad f^*=T_c(f^*+cd^*).$$

It is known that the operator T_c is Newton differentiable from $L^q\mapsto L^2$ for any $q>2$ (cf. [9]). To apply semismooth Newton algorithm, we may choose $c=\beta/\gamma$ and eliminate d^* from the optimality system:

$$\begin{cases} p^*=-M^t(Mf^*-g^\delta),\\ f^*=T_c\Big(\dfrac{1}{\gamma}p^*\Big). \end{cases} \tag{2.3}$$

Let

$$F(z)=\begin{pmatrix} p+M^t(Mf-g^\delta)\\ f-T_c\Big(\dfrac{1}{\gamma}p\Big)\end{pmatrix},$$

https://doi.org/10.4208/aamm.2014.m722 Published online by Cambridge University Press

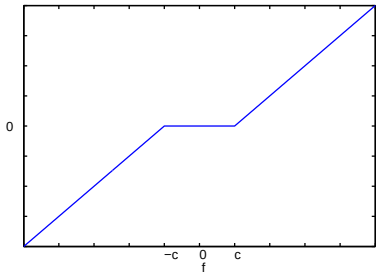


Figure 1: soft-thresholding operator.

where

$$z=\begin{pmatrix} p \\ f \end{pmatrix}.$$

Since M is a smoothing operator, due to the regularity gap between p^*/γ and f^*, $F(z)$ is Newton differentiable from any L^q to L^2 with $q>2$. One step semismooth Newton iteration reads

$$z^{k+1}-z^k=-[\nabla_N F(z^k)]^{-1}F(z^k),$$

where $\nabla_N F$ is the Newton derivative of F. In particular, let

$$A=\left\{x:\left|\frac{1}{\gamma}p(x)\right|>c\right\},\quad I=\left\{x:\left|\frac{1}{\gamma}p(x)\right|\le c\right\},$$

then $\nabla_N F\begin{pmatrix} p \\ f \end{pmatrix}$ can be chosen as

$$\nabla_N F\begin{pmatrix} p \\ f \end{pmatrix}=\begin{bmatrix} I & M^tM \\ -\frac{1}{\gamma}\chi_A & I \end{bmatrix},$$

where χ_A is the characteristic function on A.

Similar as in [9,30], semismooth Newton method can be equivalently represented by an active set approach, see Algorithm 1 for the details. The stopping condition at line 5 can be chosen as $A_\pm^k=A_\pm^{k+1}$, where

$$A_\pm^k=\left\{x:\pm\frac{1}{\gamma}p^{k-1}(x)>c\right\}.$$

It is easy to see that $A_\pm^k=A_\pm^{k+1}$ implies the convergence of the solution.

Now we will prove the locally superlinear convergence of semismooth Newton method. Denoted by $e^k=f^k-f^*$, $s^k=p^k-p^*$, subtracting Eq. (2.3) from (2.4) we obtain the error equation:

$$M^tMe^k+s^k=0, \tag{2.5a}$$

$$e^k-\frac{1}{\gamma}\chi_{A^k}s^k=T_c\left(\frac{1}{\gamma}p^{k-1}\right)-T_c\left(\frac{1}{\gamma}p^*\right)-\frac{1}{\gamma}\chi_{A^k}s^{k-1}\triangleq R^{k-1}. \tag{2.5b}$$

https://doi.org/10.4208/aamm.2014.m722 Published online by Cambridge University Press

Algorithm 1 Semismooth Newton method algorithm.

1: Given initial guess p^0, f^0.
2: **for** $k=1,2,\cdots$, **do**
3: Define the active set A^k and inactive set I^k respectively by

$$A^k=\left\{x:\left|\frac{1}{\gamma}p^{k-1}(x)\right|>c\right\},\quad I^k=\left\{x:\left|\frac{1}{\gamma}p^{k-1}(x)\right|\leq c\right\}.$$

4: Let (p^k,f^k) solves following equation:

$$\begin{cases} M^tMf^k+p^k=M^tg^\delta, \\ f^k|_{I^k}=0,\quad \left(\frac{1}{\gamma}p^k-f^k\right)\Big|_{A^k}=c\cdot sgn(p^{k-1}(x))|_{A^k}. \end{cases} \tag{2.4}$$

5: Check the stopping criterion.
6: **end for**

By noticing that χ_{A^k}/γ is Newton derivative of $T_c(p^{k-1}/\gamma)$, the remainder term R^{k-1} can be represented by

$$R^{k-1}=T_c\left(\frac{1}{\gamma}p^{k-1}\right)-T_c\left(\frac{1}{\gamma}p^*\right)-D_NT_c\left(\frac{1}{\gamma}p^{k-1}\right)(p^{k-1}-p^*),$$

and hence (see [9]):

$$\|R^{k-1}\|=o(\|p^{k-1}-p^*\|_{L^q}),\quad \forall q>2.$$

In particular, by Sobolev embedding theorem, one has

$$\|R^{k-1}\|_{L^2}=o(\|s^{k-1}\|_{H^2}). \tag{2.6}$$

Now we multiply (2.5a) with e^k and (2.5b) with s^k and subtract each other

$$\|Me^k\|^2+\frac{1}{\gamma}\|\chi_{A^k}s^k\|^2=-\langle R^{k-1},s^k\rangle. \tag{2.7}$$

Now, due to the regularity estimate (2.2) and the self-adjointness of the operator M, we have

$$\|s^k\|_{H^2}=\|M^t(Me^k)\|_{H^2}\leq C\|Me^k\|.$$

By Young's inequality we deduce that

$$\|Me^k\|\leq\delta\|s^k\|^2+C_\delta\|R^{k-1}\|^2.$$

Therefore,

$$\|s^k\|_{H^2}\leq C\|R^{k-1}\|=o(\|s^{k-1}\|_{H^2}).$$

https://doi.org/10.4208/aamm.2014.m722 Published online by Cambridge University Press

The estimate for e^k can be obtained from (2.5b):

$$\|e^k\| \leq \frac{1}{\gamma}\|s^k\| + \|R^{k-1}\|.$$

Then we have the following theorem:

Theorem 2.2. *If (p^0, f^0) closes to (p^*, f^*) enough, then (p^k, f^k) converges to (p^*, f^*) superlinearly.*

3 Discretization

In this section, we consider the numerical discretization of the semismooth Newton method. To do this, we introduce a new variable $u = Mf$, then the Newton iteration can be rewritten as

$$\begin{cases} u^k = Mf^k, \\ M^t(u^k - g^\delta) + p^k = 0, \\ f^k|_{I^k} = 0, \quad \left(\frac{1}{\gamma}p^k - f^k\right)\Big|_{A^k} = (c \cdot sgn(p^{k-1}))|_{A^k}. \end{cases} \tag{3.1}$$

One may find that each Newton iteration involves a system of differential equations. By any discretization, one needs to solve a very large linear system which involves all degree of freedom in temporal-spatial space. It can be very expensive.

Further, we notice that the operator M only involves the spacial variable and the optimal solution is supposed to be sparse. Therefore we use a different approach: Firstly, we give an explicit form of the discretization of the operator M, then use the explicit expression of discrete operator $M_{h,\tau}$ in the iterative algorithm. In principle this approach does not depend on the discretization in temporal-spatial variables. But to keep the sparse constraint, the finite element discretization is a natural choice. We will use finite element approximation in spatial variable (but the finite element spaces can be different for the source term and the final observation data) and a (weighted) finite difference approximation for fractional time derivative.

3.1 Discretization of the forward operator

Firstly we consider the discretization for the governing Eq. (1.1) with any $f \in L^2(\Omega)$. Let T_h be a quasi-uniform triangulation of Ω and V_h be the finite element spaces over T_h (for simplicity we consider continuous P_1 element), i.e.,

$$V_h = \{v : v \in C_0(\Omega),\ v|_{\Delta_h} \in P_1(\Delta_h),\ \forall \Delta_h \in T_h\}.$$

The time interval $[0,T]$ is divided into L equal subintervals by $0 = t_0 < t_1 < \cdots < t_{L-1} < t_L = T$, with $t_k = k\tau$, $\tau = T/L$, then the fraction time derivative $\frac{\partial^\alpha u(x,t)}{\partial t^\alpha}$ at t_k can be approximated

https://doi.org/10.4208/aamm.2014.m722 Published online by Cambridge University Press

by

$$\begin{aligned}\frac{\partial^\alpha u(x,t)}{\partial t^\alpha}\Big|_{t=t_k} &= \frac{1}{\Gamma(1-\alpha)}\int_0^{t_k}(t_k-\eta)^{-\alpha}\frac{\partial u(x,\eta)}{\partial \eta}\mathrm{d}\eta \\ &= \frac{1}{\Gamma(1-\alpha)}\sum_{l=1}^{k}\int_{t_{l-1}}^{t_l}(t_k-\eta)^{-\alpha}\frac{\partial u(x,\eta)}{\partial \eta}\mathrm{d}\eta \\ &\approx \frac{\tau^{-\alpha}}{\Gamma(2-\alpha)}\sum_{l=1}^{k}(u(x,t_l)-u(x,t_{l-1}))((k+1-l)^{1-\alpha}-(k-l)^{1-\alpha}) \\ &= \frac{\tau^{-\alpha}}{\Gamma(2-\alpha)}\sum_{l=1}^{k}\omega_l(u(x,t_{k+1-l})-u(x,t_{k-l})),\end{aligned}$$

where $\omega_l = l^{1-\alpha}-(l-1)^{1-\alpha}$, $l=1,\cdots,L$. For $\{u_h^k\}\subseteq V_h$, denoted by

$$Du_{h,\tau}^{\alpha,k} = \frac{\tau^{-\alpha}}{\Gamma(2-\alpha)}\sum_{l=1}^{k}\omega_l(u_h^{k+1-l}-u_h^{k-l}).$$

Therefore for any given $f\in L^2(\Omega)$, the discrete version of Eq. (1.1) reads: find $\{u_h^k\}\subseteq V_h$, such that

$$\langle Du_{h,\tau}^{\alpha,k},\psi_h\rangle+\langle\nabla u_h^k,\nabla\psi_h\rangle=\langle f,\psi_h\rangle,\quad \forall\psi_h\in V_h,\quad k=1,\cdots,L\quad\text{and}\quad u_h^0=0. \tag{3.2}$$

Next we consider the discretization to the source term f and the forward operator M. One may choose a different finite element space (either P_0 or continuous P_1 element) $W_{\tilde{h}}\subset L^2(\Omega)$ over another quasiuniform triangulation $T_{\tilde{h}}$ for f. Then the discrete forward operator $M_{h,\tau}$ is a linear transformation from $W_{\tilde{h}}$ to V_h. Let $\{\phi_j\}_{j=1}^n$ and $\{\varphi_j\}_{j=1}^m$ form a standard basis of V_h and $W_{\tilde{h}}$, respectively. For any φ_i, we can solve the Eq. (3.2) with $f=\varphi_i$ and find the representation to the corresponding solution: $M_{h,\tau}(\varphi_i)=u_h^L(\varphi_i)=\sum_{j=1}^n a_{ij}\phi_j$. Let the matrix $\Psi\in\mathbb{R}^{n\times m}$ with explicit form:

$$\Psi_{j,i}=a_{ij},\quad i=1,\cdots,m,\quad j=1,\cdots,n.$$

Therefore

$$M_{h,\tau}(\varphi_i)=\sum_{j=1}^{n}\Psi_{j,i}\phi_j\quad\text{and}\quad M_{h,\tau}^t(\phi_i)=\sum_{j=1}^{m}\Psi_{i,j}\varphi_j.$$

3.2 Discrete semi-smooth Newton method

Now we revisit the first line in Eq. (2.4): $M^tMf^k+p^k=M^tg^\delta$. The discrete version of this equation reads (we omit k for simplicity):

$$M_{h,\tau}^tM_{h,\tau}f_{h,\tau}+p_{h,\tau}=M_{h,\tau}^tg_h^\delta, \tag{3.3}$$

https://doi.org/10.4208/aamm.2014.m722 Published online by Cambridge University Press

where g_h^δ is the L^2 projection of g^δ to V_h, i.e., $g_h^\delta=\sum_{i=1}^n g_i\phi_i$, such that $\langle g_h^\delta,\phi_j\rangle=\langle g^\delta,\phi_j\rangle$ for all ϕ_j. Let $f_{h,\tau}=\sum_{i=1}^m f_i\varphi_i$, $p_{h,\tau}=\sum_{i=1}^m p_i\varphi_i$, then we have

$$\begin{cases} M_{h,\tau}^t M_{h,\tau} f_{h,\tau}=M_{h,\tau}^t M_{h,\tau}\sum\limits_{i=1}^m f_i\varphi_i=M_{h,\tau}^t\Big(\sum\limits_{i=1}^m\sum\limits_{j=1}^n f_i\Psi_{j,i}\phi_j\Big)=\sum\limits_{j=1}^n\Big(\sum\limits_{i=1}^m f_i\Psi_{j,i}\Big)M_{h,\tau}^t\phi_j \\ \qquad\qquad =\sum\limits_{j=1}^n\Big(\sum\limits_{i=1}^m f_i\Psi_{j,i}\Big)\sum\limits_{k=1}^m\Psi_{j,k}\varphi_k=\sum\limits_{k=1}^m\Big[\sum\limits_{j=1}^n\Big(\sum\limits_{i=1}^m f_i\Psi_{j,i}\Big)\Psi_{j,k}\Big]\varphi_k, \\ p_{h,\tau}=\sum\limits_{i=1}^m p_i\varphi_i, \\ M_{h,\tau}^t g_h^\delta=\sum\limits_{i=1}^n g_i M_{h,\tau}^t\phi_i=\sum\limits_{i=1}^n g_i\sum\limits_{k=1}^m\Psi_{i,k}\varphi_k=\sum\limits_{k=1}^m\Big(\sum\limits_{i=1}^n\Psi_{i,k}g_i\Big)\varphi_k. \end{cases}$$

Substitute them into Eq. (3.3) and denote by

$$\vec{f}=\begin{pmatrix} f_1 \\ \cdots \\ f_m \end{pmatrix},\quad \vec{p}=\begin{pmatrix} p_1 \\ \cdots \\ p_m \end{pmatrix},\quad \vec{g}=\begin{pmatrix} g_1 \\ \cdots \\ g_n \end{pmatrix},$$

we have the matrix form of Eq. (3.3):

$$\Psi^t\Psi\vec{f}+\vec{p}=\Psi^t\vec{g}.$$

Since $W_{\tilde{h}}$ is chosen as either P_0 or nodal P_1 element space, the active set in Algorithm 1 and the second line in Eq. (2.4) has a natural discrete version: active/inactive set is the subset of indexes for standard basis and the second line in Eq. (2.4) can be computed componentwise, see Algorithm 2 for details. The stopping condition in line 5 can be chosen as $A_\pm^k=A_\pm^{k-1}$, where $A_\pm^k=\{i:\pm\frac{1}{\gamma}p_i^{k-1}>c\}$.

One can verify that Eq. (3.5) can be rewritten as an equivalent form:

$$\begin{cases} \vec{f}^k|_{I^k}=0, \\ (\Psi_{A^k}^t\Psi_{A^k}+\gamma I_{A^k})\vec{f}^k|_{A^k}=\Psi_{A_k}^t\vec{g}+\gamma c\cdot sgn(\vec{p}^{k-1})|_{A^k}, \\ \vec{p}^k=\Psi^t(\vec{g}-\Psi\vec{f}^k). \end{cases} \tag{3.4}$$

Remark 3.1. We make several comments on this approach.

(1) The most time consuming part in this method is to find the explicit form of the discretization matrix Ψ. One needs to solve a large number of the forward problem (3.2) with $f=\varphi_i$, $i=1,\cdots,m$. Fortunately, this can be done by a parallel manner. Moreover when $n\ll m$ one can further reduce the computational cost by: letting $f=\phi_i$, $i=1,\cdots,n$ to solve the forward problem n times, then finding the L^2-projection onto V_h for each φ_i to obtain the matrix Ψ.

(2) Once the matrix Ψ is computed, we do not need to solve any partial differential equations. It is important for the inverse problem, since to pickup a good regularization parameter one needs to solve optimization problem (1.5) several times for different parameters β and γ.

https://doi.org/10.4208/aamm.2014.m722 Published online by Cambridge University Press

Algorithm 2 Discrete semismooth Newton method algorithm.

1: Given initial guess $\vec{p}^0=\begin{pmatrix} p_1^0 \\ \cdots \\ p_m^0 \end{pmatrix}$, $\vec{f}^0=\begin{pmatrix} f_1^0 \\ \cdots \\ f_m^0 \end{pmatrix}$.

2: **for** $k=1,2,\cdots$, **do**

3: Define the active set A^k and inactive set I^k respectively by

$$A^k=\left\{i:\left|\frac{1}{\gamma}p_i^{k-1}\right|>c\right\},\quad I^k=\left\{i:\left|\frac{1}{\gamma}p_i^{k-1}\right|\leq c\right\}.$$

4: Let $(\vec{p}^k,\vec{f}^k)$ solves following equation:

$$\begin{cases} \Psi^t\Psi\vec{f}^k+\vec{p}^k=\Psi^t\vec{g}, \\ f_i^k=0,\quad \forall i\in I^k,\quad \dfrac{1}{\gamma}p_i^k-f_i^k=c\cdot sgn(p_i^{k-1}),\quad \forall i\in A^k. \end{cases} \tag{3.5}$$

5: Check the stopping criterion.

6: **end for**

(3) In general, Ψ is a large nonstructural dense matrix. But during the iteration, we don't need to solve a full linear system. The linear system in second line of Eq. (3.4) is with size $|A^k|$, which is usually small.

4 Numerical examples

In this section several numerical examples in one dimensional and two dimensional domains by the proposed algorithm 2 are given. First we will give the setting for the numerical tests, which include computational domain, data generation and how to choose the initial guess and the regularization parameter. Second we use five numerical examples to show the efficiency of the algorithm.

4.1 Experiments setting

The computational domain $\Omega=(0,1)$ and $\Omega=(0,1)\times(0,1)$ for all one dimensional and two dimensional examples, respectively. The finite element space V_h (for final time measurement) and $W_{\tilde{h}}$ (for the source term) are both continuous P_1 over possible different triangulations T_h and $T_{\tilde{h}}$. In one dimensional case, $T_{\tilde{h}}$ is chosen as a uniform partition with $\tilde{h}=1/100$, T_h is also a uniform partition with either $h=1/100$ or $h=1/50$ (in Example 4.3). In the later case, the discrete matrix is 49×99, not a square matrix. In two dimensional case, we let $T_h=T_{\tilde{h}}$ with 1024 triangles. The final time is chosen as $T=1$ and the

https://doi.org/10.4208/aamm.2014.m722 Published online by Cambridge University Press

temporal discretization parameter τ is $1/100$ and $1/80$ in one or two dimensional case, respectively. To obtain the (noisy) observation data g^δ, we first give the true solution $f^\dagger$ and solve the direct problem (1.1), then add pointwise noise by $g^\delta(x_i)=u_h^L(x_i)*(1+\delta\xi)$, where x_i is the nodal point in T_h, δ is the noise level and ξ is a uniform random variable in $[-1,1]$.

There are two regularization parameters β and γ in the elastic-net regularization (1.5), where β controls the sparsity of the solution and γ improves the regularity of it. To ensure a good approximation of the elastic-net regularization, the properly chosen regularization parameters play important role. Since in the semismooth Newton algorithm, the free parameter c is chosen as β/γ, we will use the strategy as in [21] to choose parameters β and γ. Firstly we fix the ratio of two regularization parameter $\eta=\beta/\gamma$ and apply discrepancy principle to choose a solution $f(\eta)$, then apply quasi-optimality criterion [32] to choose η. In the numerical tests, we give a decreasing sequence of $\{\eta_k\}$ and choose η_k as the minimizer of $\|f(\eta_k)-f(\eta_{k-1})\|$.

For any fixed ratio $\eta=\beta/\gamma$, to obtain a solution $f(\eta)$ which satisfies the discrepancy principle, we need to solve a sequence of minimization problem with different β. Meanwhile, a good initial guess is very important for the semismooth Newton method due to its locally superlinear convergence property. Therefore we apply a continuation strategy for the regularization parameter β (cf. [6,11]), i.e., given a decreasing sequence $\{\beta_k\}$, then we solve the regularization problem with $\beta=\beta_k$ sequentially by semismooth Newton method, where the initial guess is chosen as the solution for $\beta=\beta_{k-1}$. Once the discrepancy principle $\|Mf_{\beta_k,\beta_k/\eta}-g^\delta\|\le\rho\delta$ is satisfied (for a given constant $\rho>1$), we stop the algorithm and let $f_{\beta_k,\beta_k/\eta}$ be $f(\eta)$.

4.2 Numerical tests

The first three examples are in one dimensional domain and the last two examples consider two dimensional domain.

Example 4.1. Let $\alpha=0.7$ and the exact source function to be

$$f(x)=\begin{cases}100(x-0.45), & 0.45\le x<0.5,\\ 100(0.55-x), & 0.5\le x\le 0.55,\\ 0, & \text{otherwise.}\end{cases}$$

Example 4.2. Let $\alpha=0.7$, then choose 9 points randomly from the discrete grid points and take the value as $5+10*rand(9,1)$ as the exact source term.

The reconstruction for Examples 4.1 and 4.2 can be found in Figs. 2 and 3, respectively. The noise level for Example 4.1 is chosen as $\delta=1\%,2\%,5\%$ and for Example 4.2 is chosen as 0.5%, 1%, 2%. From these two figures, we find the location of the support can be approximately found and the reconstruction is reasonable.

https://doi.org/10.4208/aamm.2014.m722 Published online by Cambridge University Press

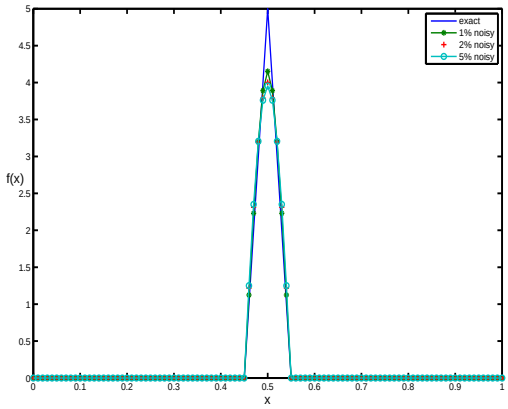


Figure 2: Reconstruction for Example 4.1.

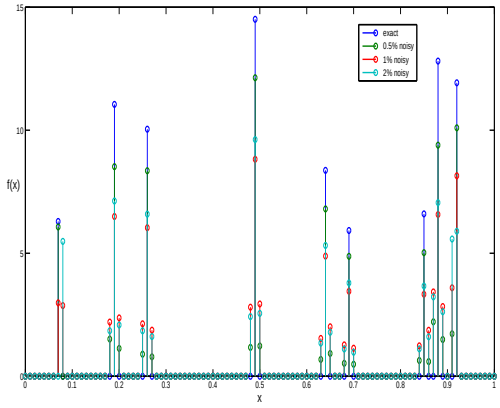


Figure 3: Reconstruction for Example 4.2.

Example 4.3. Let $\alpha=0.7$, then choose 3 points randomly from the discrete grid points and take the value as $5+10*rand(3,1)$ as the exact source term.

In this example, we consider two different discretizations for the forward time-fractional diffusion equation, i.e., $h=1/100$ and $h=1/50$. In the later case the discrete matrix is a 49×99 matrix. The noise level is also chosen as 0.5%,1% and 2%, respectively. Reconstruction by two discretizations can be found in Fig. 4.

Example 4.4. Let $\alpha=0.6$ and the exact source function to be

$$f(x)=\begin{cases} 100\Big(\dfrac{1}{100}-\Big(x-\dfrac{1}{2}\Big)^2-\Big(y-\dfrac{1}{2}\Big)^2\Big)\exp\Big[-\dfrac{(x-\frac{1}{2})^2+(y-\frac{1}{2})^2}{2}\Big], & \Big(x-\dfrac{1}{2}\Big)^2+\Big(y-\dfrac{1}{2}\Big)^2\leq\dfrac{1}{100},\\ 0, & \text{otherwise.}\end{cases}$$

Numerical reconstructions with noise level $\delta=1\%,5\%$ and the error function with $\delta=5\%$ are shown in Fig. 5. One may find the relatively error is small in this case.

https://doi.org/10.4208/aamm.2014.m722 Published online by Cambridge University Press

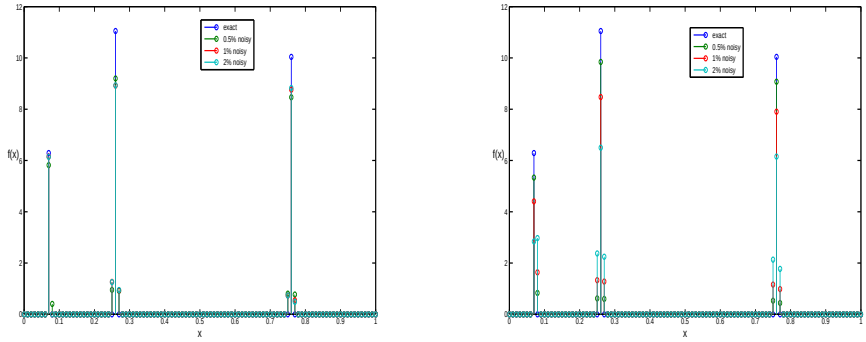
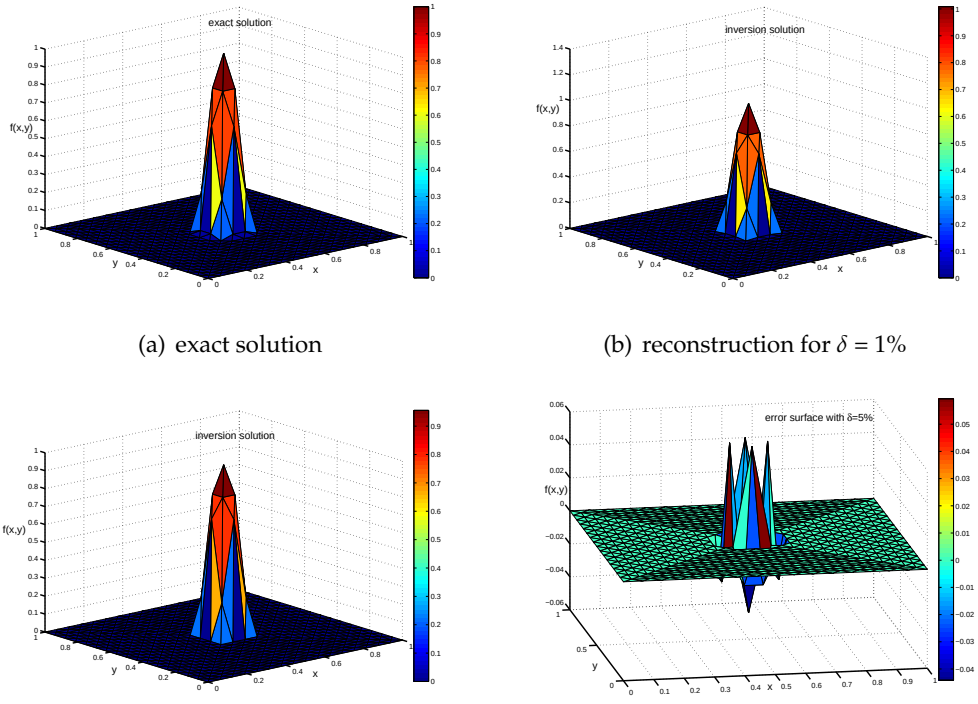


Figure 4: Comparison for different discretizations, $h=1/100$ (left panel) and $h=1/50$ (right panel).



(a) exact solution

(b) reconstruction for δ = 1%

(c) reconstruction for δ = 5%

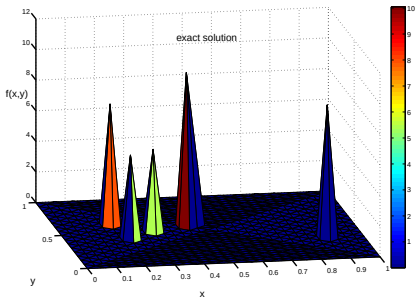
(d) error function for δ = 5%

Figure 5: Numerical Example 4.4.

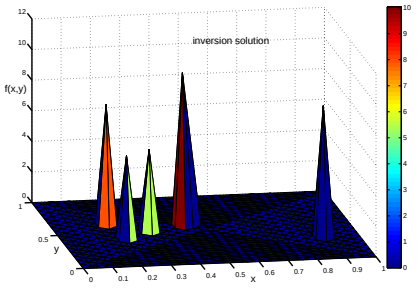
Example 4.5. Let $\alpha=0.6$, then choose 5 points randomly from the discrete grid points and take the value as $5+10*rand(5,1)$ as the exact source term.

Similar as Example 4.4, numerical results are showed in Fig. 6 with noise level $\delta=$ 1%,5% and the error function with $\delta=5\%$, respectively.

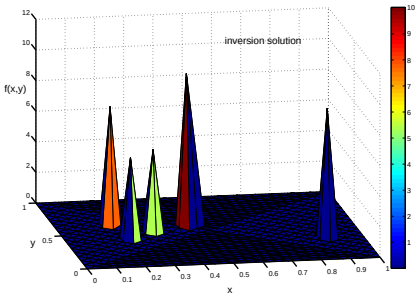
https://doi.org/10.4208/aamm.2014.m722 Published online by Cambridge University Press



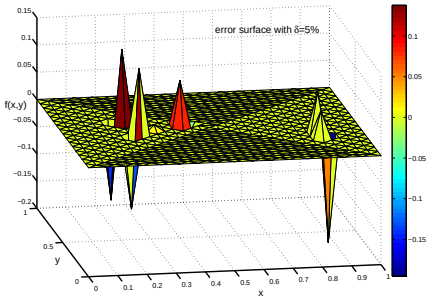
(a) exact solution



(b) reconstruction for $\delta = 1\%$



(c) reconstruction for $\delta = 5\%$



(d) error function for $\delta = 5\%$

Figure 6: Numerical Example 4.5.

5 Conclusions

In this paper, we study an inverse problem for identifying the temporally independent source function for the time-fractional diffusion equation. By elastics-net regularization, we transform the inverse problem into the optimization problem with L^1-L^2 penalties. The discrete semismooth Newton method is applied to solve the optimization problem. This approach is quite efficient by coupling with a continuation strategy on the regularization parameter. Numerical examples are given to verify the efficiency of the proposed algorithm.

Acknowledgments

The authors would like to thank the anonymous referees for their constructive comments, which have led to an improvement in the presentation of the paper. The research of J. Z. Yang is partially supported by National Science Foundation of China No. 11171305 and No. 91230203 and the work of X. Lu is partially supported by National Science Foundation of China No. 11471253, the Fundamental Research Funds for the Central Universities (13lgzd07) and the PSTNS of Zhu Jiang in Guangzhou city (2011J2200099).

https://doi.org/10.4208/aamm.2014.m722 Published online by Cambridge University Press

References

[1] J. CHENG, J. NAKAGAWA, M. YAMAMOTO AND T. YAMAZAKI, *Uniqueness in an inverse problem for a one-dimensional fractional diffusion equation,* Inverse Problems, 25(11) (2009), 115002.

[2] C. CLASON AND K. KUNISCH, *A duality-based approach to elliptic control problems in non-reflexive banach spaces,* ESAIM: Control, Optimisation and Calculus of Variations, 17(01) (2011), pp. 243–266.

[3] L. DEBNATH, *Recent applications of fractional calculus to science and engineering*, Int. Math. Math. Sci., 54 (2003), pp. 3413–3442.

[4] S. D. EIDELMAN AND A. N. KOCHUBEI, *Cauchy problem for fractional diffusion equations,* J. Differential Equations, 199(2) (2004), pp. 211–255.

[5] I. EKELAND AND R. TEMAM, Convex Analysis and Variational Problems, volume 28, SIAM, 1999.

[6] Q. FAN, Y. JIAO AND X. LU, *A primal dual active set algorithm with continuation for compressed sensing,* Signal Process. IEEE Trans., 62 (2014), pp. 6276–6285.

[7] Q. FAN, Y. JIAO, X. LU AND Z. SUN, L^q-regularization for the inverse robin problem, J. Inverse Ill-Posed Problems, (2014).

[8] R. GRIESSE AND D. A. LORENZ, *A semismooth newton method for tikhonov functionals with sparsity constraints,* Inverse Problems, 24(3) (2008), 035007.

[9] M. HINTERMÜLLER, K. ITO AND K. KUNISCH, *The primal-dual active set strategy as a semismooth newton method,* SIAM J. Optim., 13(3) (2002), pp. 865–888.

[10] Y. JIANG AND J. MA, *High-order finite element methods for time-fractional partial differential equations,* J. Comput. Appl. Math., 235(11) (2011), pp. 3285–3290.

[11] Y. JIAO, B. JIN AND X. LU, *A primal dual active set with continuation algorithm for the l^0-regularized optimization problem,* Appl. Comput. Harmonic Anal., in press, (2014).

[12] B. JIN, R. LAZAROV, J. PASCIAK AND Z. ZHOU, *Galerkin fem for fractional order parabolic equations with initial data in h^{-s}, $0<s\leq 1$,* Numer. Anal. Appl., (2013), pp. 24–37.

[13] B. JIN, D. A. LORENZ AND S. SCHIFFLER, *Elastic-net regularization: error estimates and active set methods,* Inverse Problems, 25(11) (2009), 115022.

[14] B. JIN AND W. RUNDELL, *An inverse problem for a one-dimensional time-fractional diffusion problem,* Inverse Problems, 28(7) (2012), 075010.

[15] A. A. A. KILBAS, H. M. SRIVASTAVA AND J. J. TRUJILLO, Theory and Applications of Fractional Differential Equations, Elsevier Science Limited, 2006.

[16] K. KUNISCH AND X. LU, *Optimal control for an elliptic system with convex polygonal control constraints,* IMA J. Numer. Anal., 33 (2013), pp. 875–897.

[17] K. KUNISCH AND X. LU, *Optimal control for elliptic systems with pointwise euclidean norm constraints on the controls,* Math. Program., 142(1-2) (2013), pp. 461–483.

[18] X. LI AND C. XU, *A space-time spectral method for the time fractional diffusion equation,* SIAM J. Numer. Anal., 47(3) (2009), pp. 2108–2131.

[19] Y. LIN AND C. XU, *Finite difference/ spectral approximations for the time-fractional diffusion equation*, J. Comput. Phys., 225(2) (2007), pp. 1533–1552.

[20] J. LIU AND M. YAMAMOTO, *A backward problem for the time-fractional diffusion equation,* Appl. Anal., 89(11) (2010), pp. 1769–1788.

[21] S. LU, S. V. PEREVERZEV, Y. SHAO AND U. TAUTENHAHN, *Discrepancy curves for multiparameter regularization,* J. Inverse Ill-Posed Problems, 18(6) (2010), pp. 655–676.

[22] R. METZLER AND J. KLAFTER, *Boundary value problems for fractional diffusion equations,* Phys. A Statist. Mech. Appl., 278(1) (2000), pp. 107–125.

https://doi.org/10.4208/aamm.2014.m722 Published online by Cambridge University Press

[23] J. NAKAGAWA, K. SAKAMOTO AND M. YAMAMOTO, *Overview to mathematical analysis for fractional diffusion equations-new mathematical aspects motivated by industrial collaboration,* J. Math. Industry, 2(10) (2010), pp. 99–108.
[24] R. NIGMATULLIN, *The realization of the generalized transfer equation in a medium with fractal geometry,* Phys. Status Solid B, 133(1) (1986), pp. 425–430.
[25] I. PODLUBNY, Fractional Differential Equations: An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of Their Solution and Some of Their Applications, 198, Academic Press, 1998.
[26] H. POLLARD, *The completely monotonic character of the mittag-leffler function* $e_\alpha(-x)$, Bulletin American Math. Society, 54(12) (1948), pp. 1115–1116.
[27] H. ROMAN AND P. ALEMANY, *Continuous-time random walks and the fractional diffusion equation,* J. Phys. A Math. General, 27(10) (1994), 3407.
[28] K. SAKAMOTO AND M. YAMAMOTO, *Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems,* J. Math. Anal. Appl., 382(1) (2011), pp. 426–447.
[29] G. WACHSMUTH AND D. WACHSMUTH, *Convergence and regularization results for optimal control problems with sparsity functional,* ESAIM: Control, Optimisation and Calculus of Variations, 17(3) (2011), pp. 858–886.
[30] G. STADLER, *Elliptic optimal control problems with* l^1*-control cost and applications for the placement of control devices,* Comput. Optim. Appl., 44(2) (2009), pp. 159–181.
[31] Z. SUN, Y. JIAO, B. JIN AND X. LU, *Numerical identification of a sparse robin coefficient,* Adv. Comput. Math., (2014), pp. 1–18.
[32] A. N. TIKHONOV AND V. GLASKO, *Use of the regularization method in non-linear problems,* USSR Comput. Math. Math. Phys., 5(3) (1965), pp. 93–107.
[33] J. WANG, Y. ZHOU AND T. WEI, *Two regularization methods to identify a space-dependent source for the time-fractional diffusion equation,* Appl. Numer. Math., 68 (2013), pp. 39–57.
[34] L. WANG AND J. LIU, *Data regularization for a backward time-fractional diffusion problem,* Comput. Math. Appl., 64(11) (2012), pp. 3613–3626.
[35] Z. WANG, S. QIU, Z. RUAN AND W. ZHANG, *A regularized optimization method for identifying the space-dependent source and the initial value simultaneously in a parabolic equation,* Comput. Math. Appl., 67(7) (2014), pp. 1345–1357.
[36] T. WEI AND Z. ZHANG, *Reconstruction of a time-dependent source term in a time-fractional diffusion equation,* Eng. Anal. Boundary Elements, 37(1) (2013), pp. 23–31.
[37] X. XIONG, J. WANG AND M. LI, *An optimal method for fractional heat conduction problem backward in time,* Appl. Anal., 91(4) (2012), pp. 823–840.
[38] X. YE AND C. XU, *Spectral optimization methods for the time fractional diffusion inverse problem,* Numer. Math. Theory Methods Appl., 6(3) (2013), pp. 499–519.
[39] Y. ZHANG AND X. XU, *Inverse source problem for a fractional diffusion equation,* Inverse Problems, 27(3) (2011), 035010.
[40] G. ZHENG AND T. WEI, *A new regularization method for the time fractional inverse advection-dispersion problem,* SIAM J. Numer. Anal., 49(5) (2011), pp. 1972–1990.
[41] H. ZOU AND T. HASTIE, *Regularization and variable selection via the elastic net,* J. Royal Statistical Society Series B (Statistical Methodology), 67(2) (2005), pp. 301–320.

https://doi.org/10.4208/aamm.2014.m722 Published online by Cambridge University Press