Computers and Mathematics with Applications 66 (2013) 630–638

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

Adjoint state method for fractional diffusion: Parameter identification

B. Maryshev [a], A. Cartalade [a], C. Latrille [b], M. Joelson [c], M.C. Néel [c,*]

[a] CEA-Saclay, DEN, DM2S, STMF, LATF, Bât. 454, F-91191 Gif-sur-Yvette, France
[b] CEA-Saclay, DEN, DPC, SECR, L3MR, Bât. 450, 91191 Gif-sur-Yvette, France
[c] Université d'Avignon et des Pays de Vaucluse, UMR 1114 EMMAH, 84018 Avignon Cedex, France

ARTICLE INFO

Keywords:
Sub-diffusion
Adjoint state method
Data inversion

ABSTRACT

Fractional partial differential equations provide models for sub-diffusion, among which the fractal Mobile–Immobile Model (fMIM) is often used to represent solute transport in complex media. The fMIM involves four parameters, among which we have the order of an integro-differential operator that accounts for the possibility for solutes to be sequestered during very long times. To guess fMIM parameters from experiments, an accurate method consists in optimizing an objective function that measures how much model solutions deviate from data. We show that solving an adjoint problem helps accurate computing of the gradient of such an objective function, with respect to the parameters. We illustrate the method by applying it on experimental data issued from tracing tests in porous media.

1. Introduction

Partial differential equations (p.d.e) involving fractional operators provide models for anomalous transport, in particular for sub-diffusion. At microscopic scale, sub-diffusion may correspond to non-Markovian stochastic processes with long-tailed distributions of waiting times, Meerschaert et al. [1]. On the macroscopic scale, it is associated with mean squared displacements proportional to t^{α} (α being smaller than 1), and with asymptotic behaviors showing power law decreases inconsistent with Fourier's and Fick's laws. Such behaviors are rather similar to that of solutions to fractional p.d.e such as the Fractional Fokker–Planck Equation (FFPE) (see [2]), time fractional Advection–Dispersion Equation (tfADE) and fractal Mobile–Immobile Model (fMIM) addressed by Schumer et al. [3] and Benson et al. [4].

Solutions to all those equations behave similarly at very long times, but differ significantly at finite times, especially at the beginning of experiments. Then, fMIM's solutions are similar to that of Fourier's law or the Advection–Dispersion Equation (ADE). Therefore, the fMIM is often used to simulate the spread of a tracer in media where physical arguments suggest that fluid and tracer particles may be arrested for very irregularly distributed durations, as in underground media and river flows (see [3,5,4]). To discuss fMIM's practical validity for given media, we need tools to guess the transport parameters that let its solutions represent at best experiments, as e.g. by optimizing the deviation between data and model solutions. We present a method that accurately determines the gradient of an objective function, measuring the deviation between data and model solutions. It uses an adjoint reformulation of the model itself.

We first recall fMIM's mathematical formulation and the physical meaning of its parameters, with essentials of a numerical method adapted to this model. Then, focusing on an objective function in integral form, we explain how an adjoint problem yields an alternative expression for its derivative with respect to model parameters. Then, we address a discrete

* Corresponding author. Tel.: +33 432722246.
E-mail address: mcneel@avignon.inra.fr (M.C. Néel).

0898-1221/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.camwa.2013.02.022

version of the model and of the objective function. We check the accuracy of this method, then real life examples illustrate the process to find the transport parameter of a given set of experimental data.

2. Fractional model for advection, dispersion and irregularly distributed sorption

In many porous media, tracer density P does not obey the Advection–Dispersion equation (ADE) and falls off as a power of time. Such asymptotic behavior is taken into account by the fractal Mobile–Immobile Model (fMIM) (see [3,5,4]).

2.1. The fractal Mobile–Immobile Model (fMIM)

It reads

$$\partial_t P = \partial_x \left[\partial_x D P_m - v P_m\right] + r, \tag{1}$$

where D is a diffusion coefficient and v a velocity, while r represents a source rate. As the Mobile–Immobile Model (MIM) (see [6]), the fMIM is based upon partitioning fluid as well as tracer into two phases, mobile and immobile, of densities P_m and $P - P_m$. The MIM assumes that the two phases exchange matter according to first order kinetics, equivalent to $P = P_m + Ke^{-\omega t} * P_m$, with $*$ representing Laplace convolution. In the fMIM, differently, P and P_m satisfy

$$P = \left(Id + \Lambda I_{0,+}^{1-\alpha}\right) P_m, \tag{2}$$

with, moreover, $0 < \alpha \leq 1$. Eq. (2) is equivalent to $P_m = \left(Id + \Lambda I_{0,+}^{1-\alpha}\right)^{-1} P$, the fractional integral $I_{0,+}^{1-\alpha}$ (of order $1-\alpha$) being defined by $I_{0,+}^{1-\alpha}\varphi(t) \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-t')^{-\alpha} \varphi\left(t'\right) dt'$ for $\alpha < 1$, $I_{0,+}^{0} = Id$ Samko et al. [7].

Parameter Λ has the dimension of a time to the power $\alpha - 1$, and weights the relative influence of operators Id and $I_{0,+}^{1-\alpha}$ in (2). System (1)–(2) was used by Schumer et al. [3] for the motion of dissolved solutes in complex media.

Eq. (1) is a conservation equation, associated with Fick's law applied to the mobile density, i.e. with the tracer flux equal to $-\partial_x D P_m + v P_m$. The limit case $\alpha = 1$ describes instantaneous exchanges between mobile and sorbed fractions of tracer, and setting $\Lambda = 0$ yields the Advection–Dispersion Equation (ADE). The more general case $\alpha < 1$ corresponds to heterogeneous distributions of exchange rates, that do not have any finite characteristic time. In the following, we will use the following equivalent form of (1):

$$\partial_t \left(P_m + \Lambda I_{0,+}^{1-\alpha} P_m\right) = \partial_x \left[\partial_x D P_m - v P_m\right] + r, \tag{3}$$

from which $P(x, t)$ is easily determined according to (2).

Values of α and Λ determine the late time behavior of solutions to (1)–(2), as we will immediately see.

2.2. Solutions of the fMIM

For $\alpha < 1$, solutions to (1)–(2) issued from short initial pulses fall off as $t^{-1-\alpha}$ at late times, as solutions to the fractional Fokker–Planck Equation [3,5]

$$\partial_t P = \partial_x D_{0,+}^{1-\alpha} [D^{(f)} \partial_x P - v P] + r, \tag{4}$$

here equivalent to the time-fractional ADE (tfADE) when v is supposed constant, which we assume here. The Riemann–Liouville derivative $D_{0,+}^{1-\alpha}$, that appears in (4), is defined by $D_{0,+}^{1-\alpha} = \partial_t I_{0,+}^{1-\alpha}$. Solutions to (1)–(2) are conveniently approximated by that of (4) when t becomes larger than $(3/\Lambda)^{\frac{1}{1-\alpha}}$ Schumer et al. [3]. If Λ is small, such times may fall out of the range of many experiments, especially in media of short length. At early times, differently, the solutions to (1)–(2) resemble that of the ADE (obtained by setting $\Lambda = 0$). It motivated addressing the fMIM, that proved more adapted than (4) for many natural media where disregarding the possibility for tracers to be retained in stagnant zones led to inconsistencies, as shown by Schumer et al. [3].

Since the fMIM represents mass spreading, it has a small scale version that helps understanding the physical meaning of its parameters, and provides simulation tools.

2.3. Stochastic version of (1)–(2)

As the ADE, System (1)–(2) rules the evolution of the density P of stochastic processes, that represent the limit of a wide class of random walks. They are made of successive steps, each divided into a mobile sub-step (of duration $\hat{\tau}$) that lets each tagged walker take a displacement of $v\hat{\tau} + \left(2D\hat{\tau}\right)^{1/2} N_n$, followed by an immobile period of duration $\left(\Lambda\hat{\tau}\right)^{1/\alpha} W_n$. The N_n and W_n are independent, the N_n being standard Gaussian random numbers, while the W_n are distributed by a positively valued Lévy's law of exponent $\alpha \leq 1$, and of characteristic function e^{-s^α} for $\alpha < 1$. Note that v represents the average velocity of each walker, *while it is moving*. After n such steps, each walker has spent time $T = n\hat{\tau}$ in the mobile phase, and

has finally been immobile for a duration distributed as $(\Lambda T)^{1/\alpha} W_1$, due to the stability property of Lévy's laws. Except for $\alpha = 1$, parameter Λ that has the dimension of a time to the power $\alpha - 1$, is just a scale linking immobile periods to random variables as W_1. In the limit case $\alpha = 1$, Λ is the time spent immobile per time spent moving. Moreover, for all $\alpha \leq 1$, walkers' trajectories tend to the sample paths of a process whose density P satisfies system (1)–(2), when time scale $\hat{\tau}$ tends to zero (see [4,5]).

The limit process $x(t)$ is a subordinated Brownian motion (see [8]) obtained from standard Brownian motion $B(t)$ (with drift) by applying a random time change $Z(t)$. It results in $x(t) = x_0 + vZ(t) + \sqrt{2D}B(Z(t))$ in the particular case where all tracer particles start from x_0 at time $t = 0$ (i.e. $r(t) = \delta(t)\delta(x - x_0)$). Moreover, for each t, the random variable $Z(t)$ is such that $Z(t) + (\Lambda Z(t))^{1/\alpha} W_1$ is distributed as number t. Parameter α weights the heterogeneity of immobile periods: smaller values of α correspond to more heterogeneous distributions, and the limit $\alpha = 1$ corresponds to trapping durations of equal value $\Lambda\hat{\tau}$.

Particle tracking using such random walks, helps representing solutions of the fMIM [4,5], beside discretization methods using approximations to $D^{\alpha}_{0,+}$ or $I^{1-\alpha}_{0,+}$ [9].

2.4. Numerical simulation of the fMIM

Solutions of (3) can be computed by approximating fractional integral $I^{1-\alpha}_{0,+}u(ih, n\tau)$ at order τ^2 by $\Sigma^n_{j=0} I(j, n)u(ih, (n-j)\tau)$. Weights $I(j, n)$, derived from trapezoidal rule [9] satisfy $I_0 = I(0, k) = \frac{\tau^{1-\alpha}}{\Gamma(3-\alpha)}$ for $k > 0$, $I(0, 0) = 0$, $I(k, k) = I_0\left[(2-\alpha)k^{1-\alpha} - k^{2-\alpha} + (k-1)^{2-\alpha}\right]$ and

$$I(j, k) = I_0\left[(j+1)^{2-\alpha} - 2j^{2-\alpha} + (j-1)^{2-\alpha}\right], \quad j = 1 \ldots k-1. \tag{5}$$

Riemann–Liouville derivatives $\partial_t I^{1-\alpha}_{0,+} = D^{\alpha}_{0,+}$ are approximated (among other possibilities) accordingly by

$$D^{\beta}_{0,+}u(mh, k\tau) \approx \frac{1}{\tau}\left(\sum_{j=0}^{k} I(j, k)\, u^{k-j}_m - \sum_{j=0}^{k-1} I(j-1, k-1)\, u^{k-j}_m\right). \tag{6}$$

We note $u(x = mh, t = k\tau) = u^k_m$ where k and m belong to $\{0 \ldots N\}$ and $\{0 \ldots N_{sp}\}$, and set $P_m = u$ for simplicity.

Inserting these approximations into (3), and using centered finite differences for space derivatives within a semi-implicit time-stepping, avoids instabilities and results in

$$\left[\frac{\tau v}{2h} - \frac{\tau D}{h^2}\right]u^k_{m+1} + \left[1 + \Lambda I_0 + \frac{2\tau D}{h^2}\right]u^k_m + \left[-\frac{\tau v}{2h} - \frac{\tau D}{h^2}\right]u^k_{m-1}$$
$$= u^{k-1}_m + \Lambda\left[\sum_{j=0}^{k-1} I(j, k-1)\, u^{k-j-1}_m - \sum_{j=1}^{k} I(j, k)\, u^{k-j}_m\right] \tag{7}$$

for $0 < m < N_{sp} - 1$. Moreover, the source term is $r = V_0 C_0 H\left(t - t_f\right)\delta(x)$, to represent uniform injection of a tracer solution (concentration C_0) at a rate of V_0 during a time t_f at point $x = 0$. At both ends we have

$$u^k_0 = \frac{\left(Du^k_1 + hV_0C_0\right)}{hv + D}, \quad k \leq \frac{t_f}{\tau}, \qquad u^k_0 = \frac{Du^k_1}{hv + D}, \quad k > \frac{t_f}{\tau} \tag{8}$$

and

$$u^k_{N_{sp}+1} = u^k_{N_{sp}}. \tag{9}$$

Implicit schemes (as above) using the weights of Diethelm et al. [9] for fractional operators were used by Maryshev et al. [10] and Zoia et al. [11]. The accuracy of the discrete version (7)–(9) of (1)–(2) was checked using exact solutions, corresponding to source terms r representing well chosen initial data, of course different from $V_0C_0H(t - t_f)\delta(x)$. With $N_{sp} = 300$ and $v\Delta t < 4\Delta x$, for example, relative errors of less than $1/1000$ were observed.

The discrete version (7)–(9) of (1)–(2) thus allows us to approximate the solutions issued from a given set (D, v, Λ, α) of parameters. Suppose now, we are given records of the tracer density P^*. To find the parameters (D, v, Λ, α) corresponding to the solutions P of (1)–(2) that fit at best P^*, we compute $\|P - P^*\|$ for several quadruples (D, v, Λ, α), searching the one that minimizes this norm. To this end, we need the gradient of $\|P - P^*\|^2$.

3. Adjoint state method for automatic differentiation

This gradient has four components. They are integrals in some space–time domain, of functions linked to P and P^*, and adjoint operators help transforming integrals into more easily computed expressions. We will detail this point with p.d.e. such as (3), before passing to a discrete version.

3.1. Problem

Our final aim is determining D, v, Λ, α such that $P = (Id + \Lambda I_{0,+}^{1-\alpha})u$ computed from the solution u of (3) deviates from data P^* as little as possible. This is equivalent to searching the minimum of functional $E(D, v, \Lambda, \alpha) = \int_0^T \int_\Omega f(D, v, \Lambda, \alpha, u(x, t))dxdt$, with f given by

$$f(x, t) = \left|\left(Id + \Lambda I_{0,+}^{1-\alpha}\right) u - P^*\right|^2. \tag{10}$$

Automatized optimization methods to find such a minimum need differentiating E with respect to parameters D, v, Λ, α. Indeed, they use solutions u to (3), computed from successive trial parameter values. After each of them, parameter values to be used at the next step need to be determined, which is done on the basis of the gradient of E. Using finite differences for that would require solving Eq. (3) eight times at each step, while the adjoint state method requires solving a similar (yet different) p.d.e. one time. The principles of the method are not new (see [12]). Yet, as far as we know, it was never applied to p.d.es that involve fractional integrals or derivatives whose order (here α) requires being identified.

3.2. Determining the gradient of functional E

With s representing D, v, Λ or α, gradient's components are the

$$\frac{\partial E}{\partial s} = \int_0^T \int_\Omega \left(\frac{\partial f}{\partial s} + \frac{\partial f}{\partial u}\frac{\partial u}{\partial s}\right) dxdt. \tag{11}$$

Instead of determining $\frac{\partial u}{\partial s}$ (which would necessitate solving (3) many times and accumulate unnecessary errors), the adjoint state method (see the book by Chavent [12]) writes $\int_0^T \int_\Omega \frac{\partial f}{\partial u}\frac{\partial u}{\partial s}dxdt$ as the scalar product of two other functions, which is easier to compute numerically. It uses the adjoint A^* in $X = L^2(\Omega \times [0, T])$ of operator A defined by

$$Au \equiv \partial_t \left(u + \Lambda I_{0,+}^{1-\alpha} u\right) - \partial_x \left[\partial_x Du - vu\right], \tag{12}$$

that appears in (3) associated to initial and boundary data $u(x, t) = 0$, $(D\partial_x u - vu)(0, t) = 0$ and $\partial_x u(1, t) = 0$. Indeed, $\frac{\partial u}{\partial s}$ satisfies

$$A\frac{\partial u}{\partial s} + \frac{\partial A}{\partial s}u = \frac{\partial r}{\partial s}u, \tag{13}$$

which implies $\int_0^T \int_\Omega \psi \cdot \frac{\partial A}{\partial s}udxdt = \int_0^T \int_\Omega \psi \frac{\partial r}{\partial s}udxdt - \int_0^T \int_\Omega \psi A\frac{\partial u}{\partial s}dxdt$ for each function ψ in X. Hence, function ψ satisfying

$$A^*\psi = -\frac{\partial f}{\partial u} \tag{14}$$

also satisfies

$$-\int_0^T \int_\Omega \psi \cdot \frac{\partial r}{\partial s}udxdt + \int_0^T \int_\Omega \psi \cdot \frac{\partial A}{\partial s}udxdt = \int_0^T \int_\Omega \frac{\partial f}{\partial u}\frac{\partial u}{\partial s}dxdt, \tag{15}$$

from which we deduce

$$\frac{\partial E}{\partial s} = \int_0^T \int_\Omega \left(\frac{\partial f}{\partial s} + \psi\frac{\partial A}{\partial s}u\right) dxdt - \int_0^T \int_\Omega \psi \cdot \frac{\partial r}{\partial s}udxdt. \tag{16}$$

Note that multiplying (13) by ψ, then integrating over $\Omega \times [0, T]$ yields an expression that balances the integral of $\frac{\partial f}{\partial u}\frac{\partial u}{\partial s}$ in (11): when we add it to this later, $\frac{\partial u}{\partial s}$ disappears. It is exactly what we will do when determining the gradient of the discrete version of our problem.

Either when dealing with the continuous or with the discrete version, we need A^*, $\frac{\partial A}{\partial s}$ and $\frac{\partial f}{\partial s}$ that involve fractional operators.

3.3. The adjoint operator of A, $\frac{\partial f}{\partial u}$, and derivatives of A

The adjoint operator of $I_{0,+}^{1-\alpha}$ is $\left(I_{0,+}^{1-\alpha}\right)^* = I_{T,-}^{1-\alpha}$ (see [7]), defined by $I_{T,-}^{1-\alpha} = \frac{1}{\Gamma(1-\alpha)} \int_t^T \left(t' - t\right)^{-\alpha} \varphi\left(t'\right) dt'$. Hence, that of operator A is

$$A^*\psi = -\partial_t\psi - \Lambda\partial_t I_{T,-}^{1-\alpha}\psi - D\partial_{x^2}^2\psi - v\partial_x\psi,$$

associated with time–space boundary data

$$\psi(x, T) = 0, \qquad D\partial_x\psi + v\psi = (1, t) = 0, \qquad \partial_x\psi(0, t) = 0. \tag{17}$$

Derivatives of f with respect to u and of A with respect to D, v, Λ are obvious, and that corresponding to parameter α is

$$-\frac{\partial A}{\partial \alpha}u = \Lambda \frac{\Gamma'(1-\alpha)}{\Gamma(1-\alpha)}\partial_t I^{\alpha}_{0,+}u + \frac{\Lambda}{\Gamma(1-\alpha)}\partial_t\left(t^{-\alpha} Ln(t) * u\right),$$

where digamma function Γ'/Γ satisfies

$$\frac{\Gamma'(z+1)}{\Gamma(z+1)} = -\gamma + \Sigma_{k=1}^{\infty}\frac{z}{k(z+k)}, \tag{18}$$

γ representing Euler's constant.

In general, instead of (3) we solve a discrete version that is an algebraic equation in a finite dimensional linear manifold and, moreover, data also are in discrete form, so that the objective function is defined by finite sums instead of integrals as in (11).

4. Discrete version

In real life, the objective function is a discrete sum instead of an integral, and we solve a finite dimensional problem that only approximates $Au = r$.

4.1. The forward problem

The discrete version of this equation is system (7)–(9), that lets pass from time step $k-1$ to time step k. It writes

$$G\cdot \mathbf{u}^k - \sum_{j=1}^{k} W(j,k)\,\mathbf{u}^{k-j} = \mathbf{F}^k \tag{19}$$

in matrix form, with $\mathbf{u}^k = \left(u_1^k, u_2^k, \ldots, u_{N_{sp}}^k\right)^t$. The discrete right hand side is $\mathbf{F}^k = \left(F_1^k, F_2^k, \ldots, F_{N_{sp}}^k\right)^t$ with $F_1^k = \frac{\tau V_0 C_0}{2h}\frac{2D+vh}{vh+D}$ for $k \le \frac{t_f}{\tau}$, and $F_m^k = 0$, for $m > 1$. Coefficients $W(j,k)$ satisfy $W(1,1) = 1 - \Lambda I_0(1-\alpha)$

$$W(1,k) = 1 + \Lambda I_0\left(3 - 2^{2-\alpha}\right), \quad k > 1, \tag{20}$$

and $W(j,k) = \Lambda\left(I(j-1,k-1) - \Lambda I(j,k)\right)$ for $j = 2\ldots k$, $k > 1$. The square matrix $G = (g_{l,s})$ in (19) is tridiagonal

$$G = \begin{pmatrix} GSL & GR & 0 & \cdots & 0 \\ GL & GM & GR & \cdots & 0 \\ 0 & GL & GM & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & GL & GSR \end{pmatrix}, \tag{21}$$

with coefficients

$$GSL = g_{1,1} = 1 + \Lambda I_0 + \frac{\tau D v}{2h(vh+D)} + \frac{\tau D}{h^2}, \qquad GL = g_{l,l-1} = -\frac{\tau v}{2h} - \frac{\tau D}{h^2}, \quad l > 1$$

$$GSR = g_{N_{sp},N_{sp}} = 1 + \Lambda I_0 + \frac{\tau D}{h} + \frac{\tau v}{2h}, \qquad GM = g_{l,l} = 1 + \Lambda I_0 + \frac{2\tau D}{h^2}, \quad l > 1$$

$$GR = g_{l,l+1} = \frac{\tau v}{2h} - \frac{\tau D}{h^2}, \quad 0 < l < N_{sp}. \tag{22}$$

4.2. Discrete objective function

The discrete objective function, that we still denote by E, is $E = \sum_{i,n} f_i^n$, with

$$f_i^n = \left[u_q^p + \Lambda\sum_{j=0}^{p}\left(I(j,p)\,u_q^{p-j,}\right) - \left(P^*\right)_i^n\right]^2,$$

$p = \frac{n\Delta t}{\tau}$, $q = \frac{i\Delta x}{h}$. Hence, the gradient of E has components

$$\frac{\partial E}{\partial s} = \sum_{i,n}\left(\sum_{k=1}^{N}\sum_{m=1}^{N_{sp}}\frac{\partial f_i^n}{\partial u_m^k}\frac{\partial u_m^k}{\partial s} + \frac{\partial f_i^n}{\partial s}\right), \tag{23}$$

and we focus on each of them, i.e. on each parameter s among D, v, Λ and α. We aim at finding some $\psi = \left(\psi_m^k, m = 1\ldots N_{sp},\ k = 1\ldots N\right)$ that helps us compute $\frac{\partial E}{\partial s}$. Inspired by Section 3.2, we search ψ such that its product by the derivative of (19) w.r.t. s equilibrates all items containing the $\frac{\partial u_m^k}{\partial s}$ in the discrete sum (23). Moreover, it must satisfy the discrete version

of the time–space boundary data (17) of the adjoint problem. For each (m, k), the $\frac{\partial u_m^k}{\partial s}$ solve the following equation, obtained by differentiating (19):

$$\sum_{l=1}^{N_{sp}} \left(\frac{\partial g_{l,m}}{\partial s} u_l^k + g_{l,m} \frac{\partial u_l^k}{\partial s} \right) - \sum_{j=1}^{k} \left(\frac{\partial W(j,k)}{\partial s} u_m^{k-j} + W(j,k) \frac{\partial u_m^{k-j}}{\partial s} \right) = \frac{\partial F_m^k}{\partial s}. \tag{24}$$

Multiplying it by ψ_m^k and summing over k and m, we obtain

$$\begin{aligned} 0 = & \sum_{k=1}^{N} \sum_{m=1}^{N_{sp}} \psi_m^k \left[\sum_{l=1}^{N_{sp}} \frac{\partial g_{l,m}}{\partial s} u_l^k - \sum_{j=1}^{k} \frac{\partial W(j,k)}{\partial s} u_m^{k-j} \right] + \sum_{k=1}^{N-1} \sum_{m=1}^{N_{sp}} \frac{\partial u_m^k}{\partial s} \\ & \times \left[\sum_{l=1}^{N_{sp}} \psi_l^k g_{m,l} - \sum_{j=1}^{N-k} \psi_m^{j+k} W(j, k+j) \right] + \sum_{m=1}^{N_{sp}} \left[\frac{\partial u_l^N}{\partial s} \sum_{l=1}^{N_{sp}} \psi_N^l g_{m,l} - \sum_{k=1}^{N} \psi_m^k \frac{\partial F_m^k}{\partial s} \right]. \end{aligned} \tag{25}$$

Adding this latter equation to (23) yields

$$\begin{aligned} \frac{\partial E}{\partial s} = & \sum_{m=1}^{N_{sp}} \frac{\partial u_l^N}{\partial s} u_l^N \left[\sum_{l=1}^{N_{sp}} \psi_l^N g_{s,l} + \sum_{i,n} \frac{\partial f_i^n}{\partial u_m^N} \right] + \sum_{k=1}^{N-1} \sum_{m=1}^{N_{sp}} \frac{\partial u_s^k}{\partial s} \left[\sum_{l=1}^{N} \psi_l^k g_{m,l} - \sum_{j=1}^{N-k} \psi_m^{k+j} W(j, k+j) \right. \\ & \left. + \sum_{i,n} \frac{\partial f_i^n}{\partial u_m^k} \right] + \sum_{k=1}^{N} \sum_{m=1}^{N_{sp}} \psi_m^k \left(\sum_{l=1}^{N_{sp}} \frac{\partial g_{l,s}}{\partial s} u_l^k - \sum_{j=1}^{k} \frac{\partial W(j,k)}{\partial s} u_m^{k-j} \right) + \sum_{i,n} \frac{\partial f_i^n}{\partial s} - \frac{\partial F_m^k}{\partial s}. \end{aligned} \tag{26}$$

4.3. Discrete adjoint state

All the $\frac{\partial u_l^N}{\partial s}$ disappear from (26) if ψ satisfies

$$\begin{aligned} \sum_{l=1}^{N_{sp}} \psi_l^N g_{m,l} &= - \sum_{i,n} \frac{\partial f_i^n}{\partial u_m^N} \\ \sum_{l=1}^{N_{sp}} \psi_l^k g_{m,l} &= \sum_{j=1}^{N-k} \psi_m^{j+k} W(j, k+j) - \sum_{i,n} \frac{\partial f_i^n}{\partial u_m^k} \end{aligned} \tag{27}$$

for $k < N$, that corresponds to Eq. (14). This requirement defines the discrete adjoint state ψ. It satisfies an evolution equation, that we must solve upon beginning with $\psi^{l,N}$, i.e. from end time T, as for the continuous version where operator $\left(I_{0,+}^{1-\alpha}\right)^*$ is $\left(I_{T,-}^{1-\alpha}\right)$. Moreover, it involves transpose matrix G^t. For derivatives $\frac{\partial f_i^n}{\partial s}$ and $\frac{\partial W}{\partial s}$, we need to use $\frac{\partial I}{\partial \alpha}(j,k) = \frac{\partial I_0}{\partial \alpha}[\ ] + I_0 \frac{\partial [\ ]}{\partial \alpha}$, with $[\ ] = (j+1)^{2-\alpha} - 2j^{2-\alpha} + (j-1)^{2-\alpha}$. In $\frac{\partial I_0}{\partial \alpha} = -\frac{\tau^{-\alpha} L n \tau}{\Gamma(3-\alpha)} + \frac{\tau^{1-\alpha} \Gamma'(3-\alpha)}{\Gamma^2(3-\alpha)}$ we have digamma function Γ'/Γ (see Eq. (18)).

Eqs. (23) and (27) imply $\frac{\partial E}{\partial s} =$

$$\sum_{k=1}^{N} \sum_{m=1}^{N_{sp}} \psi_m^k \left(\sum_{l=1}^{N_{sp}} \frac{\partial g_{l,m}}{\partial s} u_l^k - \sum_{j=1}^{k} \frac{\partial W(j,k)}{\partial s} u_m^{k-j} - \frac{\partial F_m^k}{\partial s} \right) + \sum_{i,n} \frac{\partial f_i^n}{\partial s}. \tag{28}$$

Once ψ has been determined by means of (27), the above equation yields all components of the gradient of E.

5. Computing the gradient of E

Let us check separately that thus computed gradients are accurate, and inserting them in an optimization process yields an efficient inversion method.

5.1. Accuracy of the discrete gradient of E

With arbitrary data, the first checks consist of solving (19) for given parameter values (D, v, Λ, α), then obtaining $\frac{\partial E}{\partial s}$ using (27)–(28). The result is then compared with finite differences, approximating $\frac{\partial E}{\partial s}$ by $\frac{E(s+ds)-E(s-ds)}{2ds}$, the $E(s \pm ds)$ being computed by solving (19) with $s \pm ds$ instead of s. Provided step ds is taken small enough, both methods return very similar approximations to $\frac{\partial E}{\partial s}$, with relative deviations of 10^{-5} in L^2 norm. The space and time steps (see Section 2.4) that provided us the relative accuracy of 1/1000 for the solutions to (3) were used. Such checks do not compare exact and computed gradients: they just compare two methods that compute the gradient of an approximation to E, both being based on the same approximation to u.

Table 1
Some steps of an optimization process in parameter space (D, v, Λ, α), tending to minimize the deviation $\sqrt{E}$ between solutions to (1), and sparse numerical data corresponding to parameters $(D^*, v^*, \Lambda^*, \alpha^*) = (0.01, 0.46, 0.84, 0.7)$ in domain $[0, L]$ with $L = 80$. At the end of the process, parameters $D^*, v^*, \Lambda^*, \alpha^*$ are retrieved with relative errors $\frac{|D-D^*|}{D^*} = 6 \times 10^{-6}$, $\frac{|v-v^*|}{v^*} = 2 \times 10^{-4}$, $\frac{|\Lambda-\Lambda^*|}{\Lambda^*} = 3 \times 10^{-4}$ and $\frac{|\alpha-\alpha^*|}{\alpha^*} = 5 \times 10^{-5}$.

i	$10^2 D$	Λ	α	v	E	$\frac{\sqrt{E}}{\|P^*\|}$
0	1.5	1.0	0.5	1.0	8.6	0.3
21	1.04	2.73	0.78	0.97	3×10^{-3}	5.5×10^{-3}
41	9.7	1.05	0.72	0.52	3×10^{-4}	1.7×10^{-3}
51	1.0	0.84	0.7	0.46	4×10^{-8}	2×10^{-5}

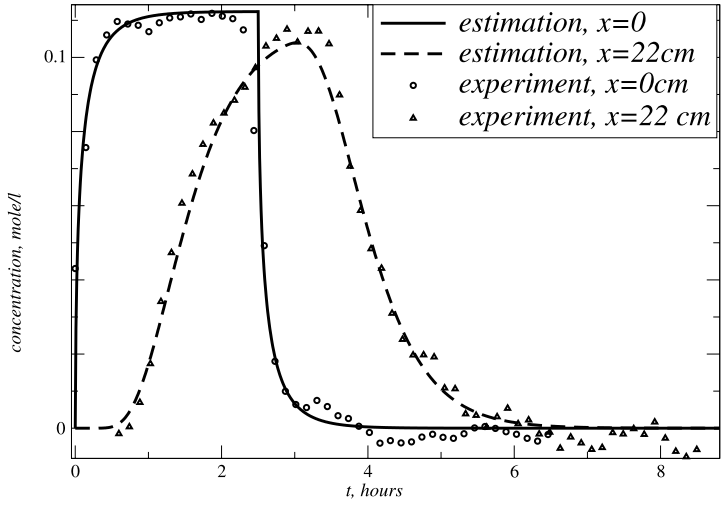

Fig. 1. Experimental data and profiles, computed from the fMIM using parameters, estimated from the data by means of the adjoint state method. A solution of concentration $C_0 = 0.11$ mole/l had been injected at a rate of $V_0 = 2.12$ cm/h during 2.5 h. Starting parameters are $D_0 = 155$ cm^2/h, $v_0 = 11.9$ cm/h, $\alpha_0 = 1$ and $\Lambda_0 = 4 \times 10^{-3}$ h$^{1-\alpha}$. Estimated parameters are $D_e = 22.27$ cm^2/h, $v_e = 12.7$ cm/h, $\alpha_e = 0.43$ and $\Lambda_e = 12.1 \times 10^{-3}$ h$^{1-\alpha}$.

5.2. Testing the synergy between gradients and optimization process

For arbitrary parameter values $(D^*, v^*, \Lambda^*, \alpha^*)$, we solve (19) to obtain a series of discrete curves $(P^*)_i^n$, and we try to retrieve the "true" parameters $(D^*, v^*, \Lambda^*, \alpha^*)$ using a simple descent method started from $(D_0, v_0, \Lambda_0, \alpha_0)$.

After a reasonable number of gradient steps we pass from initial guess $(D_0, v_0, \Lambda_0, \alpha_0)$ represented on the first line of Table 1 to a very accurate estimation of $(D^*, v^*, \Lambda^*, \alpha^*)$ corresponding to a relative error $\sqrt{E}/\|P^*\|$ of about 10^{-5}, provided the process is started not too far from the "true" parameters. For instance, the optimization process represented in Table 1 corresponds to 51 steps and takes one hour, using a laptop.

6. Processing experimental data

The above presented adjoint state method was applied to data, in the form of time-concentration (discrete) curves, recorded in non-saturated sand columns by means of X-ray spectrometry. It is a non-invasive method that allows to locally measure water content θ as well as total concentration C inside a porous column where, moreover, a steady water flow is imposed. The latter condition is necessary (not sufficient) to have a steady medium, especially regarding the water content.

6.1. Experiment and experimental device

In a column filled of sand, appropriate application of suction results in steady water flow rate V and water content θ, V being measured at the inlet and outlet. Tracer concentration and water content are measured locally on the basis of their influence on the intensity of X-ray beams traversing cross sections of the column. Different wavelengths (adapted to the medium) are used to measure water content θ and tracer concentration C^*. Moreover, the X-ray source and sensor can be moved along the column to record each desired quantity, at arbitrary locations successively. We pass to the tracer density P^* according to $P^* = \theta C^*$. Repetitions confirm that the oscillations observed on the data represented on Figs. 1–2 *are* noise, which is also suggested by the negative values observed on profile tailings. Indeed, the concentration cannot be negative. Yet, this disadvantage is the price to pay for non-invasive measurements of tracer concentration *inside* the column.

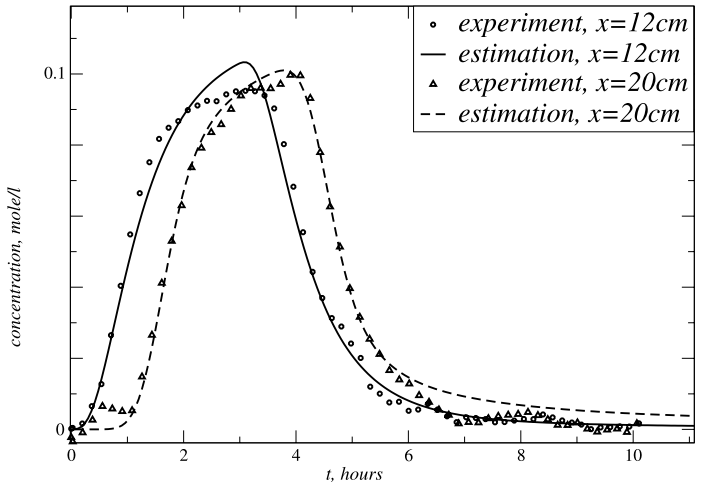

Fig. 2. Experimental data and profiles, computed from the fMIM using parameters, estimated from the data by means of the adjoint state method. A solution of concentration $C_0 = 0.1$ mole/l had been injected at a rate of $V_0 = 1.05$ cm/h for 3 h. Starting parameters are $D_0 = 20$ cm^2/h, $v_0 = 1.05$ cm/h, $\alpha_0 = 0.5\Lambda_0 = \times 10^{-3}$ h$^{1-\alpha}$. Estimated parameters are $D_e = 0.482$ cm^2/h, $v_e = 10.54$ cm/h, $\alpha_e = 0.893\Lambda_e = 5.85 \times 10^{-3}$ h$^{1-\alpha}$.

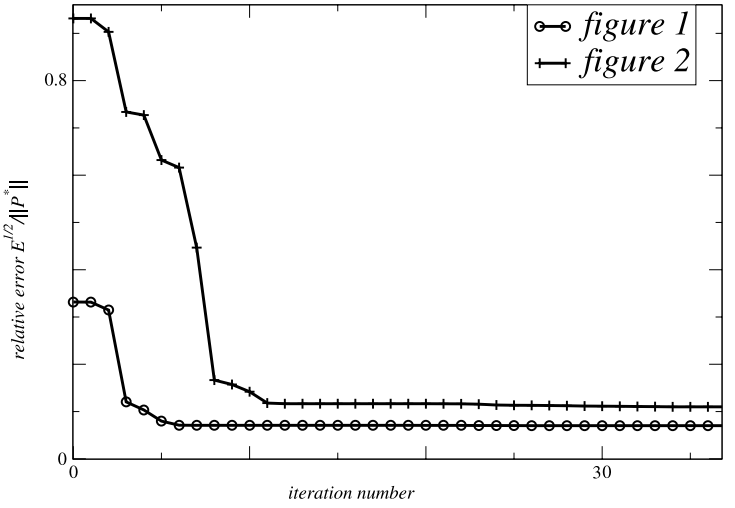

Fig. 3. Relative deviations between data and concentration profiles deduced from the inspected parameters along the optimization process, for the data of Figs. 1 and 2.

6.2. Experimental data

Two examples of the resulting time-concentration profiles are represented by symbols on the figures. In both cases, step shaped input tracer pulses were applied at the inlet. Unfortunately, more clever inputs used for signal processing (as e.g. by Sabatier et al. [13]) cannot be achieved easily for tracer fluxes. Of course, we checked the step-shaped input flux at column's inlet (using electrical conductivity).

Figs. 1–2 correspond to two sets of data, defining two objective functions. In both cases, the optimization process was started from initial values $(D_0, v_0, \Lambda_0, \alpha_0)$ suggested by commonly used physical approximations for D_0 and v_0. Any value of α_0 between 0 an 1 can be used, and exploratory attempts suggest using a small value for Λ_0 (for both data sets). Examples of such initial values are detailed in figure captions. Recording the objective function during the optimization process provided the relative deviation between the data and the concentration profiles deduced from the parameters inspected at each step (see Fig. 3). The process is stopped when it is found stationary, at parameter values $(D_e, v_e, \Lambda_e, \alpha_e)$: profiles deduced from $(D_e, v_e, \Lambda_e, \alpha_e)$ are represented by lines on the figures. They show relative deviations from data of 0.07 for Fig. 1 and 0.11 for Fig. 2.

A Laplace transform based method (see [14]) was also applied on the same data. It was found accurate (though less than the present one) when checked by using synthetic data, provided observations times were not too short. Moreover, it does not adapt to concentration records showing negative values (irrelevant for the concentration). In the counterpart, it does not involve any iteration and is very rapid. When it was applied to the data of Fig. 1, it could not find any admissible value for α.

Yet, it found estimates of the other parameters, not very different from (D_e, v_e, Λ_e), with a relative error of 0.076, insensitive to the variations of α. For the data of Fig. 2, it provided estimates of (D, v) not far from (D_e, v_e), and non-admissible estimates for Λ and α. The issues of this method provide useful D_0 and v_0.

7. Conclusion

The adjoint state method is a powerful auxiliary tool for inversion methods that aim at determining the parameters of a partial differential equation that minimize an objective function. It was extended to fractional p.d.e, a class of equations involving derivatives whose order often needs to be guessed. In a first step, we applied it to problems with parameters independent of variables x and t. Yet, a significant advantage of this method is that it extends to the case when this condition ceases to be satisfied.

The comparisons presented here between model (1)–(2) and data are not as satisfactory as desired. This is partly due to significant uncertainties caused by the method that measures the tracer concentration inside the porous media. More importantly, water content records suggest that the parameters depend on x: currently ongoing work adapting the adjoint state method to x dependent parameters confirms that relaxing this assumption significantly improves the agreement between model and experiment.

References

[1] M.M. Meerschaert, H.P. Scheffler, J. Appl. Probab. 41 (2) (2004) 455.
[2] R. Metzler, J. Klafter, Phys. Rep. 339 (2000) 1.
[3] R. Schumer, D.A. Benson, M.M. Meerschaert, B. Bauemer, Water Resour. Res. 39 (10) (2003) 1296.
[4] D.A. Benson, M.M. Meerschaert, Adv. Water Resour. 32 (2009) 532.
[5] Y. Zhang, D.A. Benson, B. Bauemer, Water Resour. Res. 44 (2008) W04424.
[6] M.T. Van Genuchten, P.J. Wierenga, Soil Sci. Soc. Am. J. 40 (4) (1976) 473.
[7] S.G. Samko, A.A. Kilbas, O.I. Marichev, Fractional Integrals and Derivatives: Theory and Applications, Gordon and Breach, New York, 1993.
[8] R. Gorenflo, F. Mainardi, A. Vivoli, Chaos Solitons Fractals 34 (2007) 87.
[9] K. Diethelm, N.J. Ford, A.D. Freed, Y. Luchko, Comput. Methods Appl. Mech. Engrg. 194 (2005) 543.
[10] B. Maryshev, M. Joelson, D. Lyubimov, T. Lyubimova, M.C. Néel, J. Phys. A: Math. Theor. 42 (2009) 115001.
[11] A. Zoia, M.C. Néel, A. Cortis, Phys. Rev. E 81 (2010) 031104.
[12] G. Chavent, Non Linear Least Squares for Inverse Problems, in: Scientific Computation, Springer, Dordrecht, Heidelberg, London, New York, 2009.
[13] J. Sabatier, M. Aoun, A. Oustaloup, G. Grégoire, F. Ragot, P. Roy, Signal Process. 86 (2006) 2645–2657.
[14] M. Ouloin, B. Maryshev, M. Joelson, C. Latrille, M.C. Néel, Transp. Porous Media, http://dx.doi.org/10.1007/S11242-012-0104-z.