INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 20, Number 3, Pages 353–370

doi: 10.4208/ijnam2023-1014

NONLOCAL-IN-TIME DYNAMICS AND CROSSOVER OF DIFFUSIVE REGIMES

QIANG DU AND ZHI ZHOU

Abstract. The aim of this paper is to study a simple nonlocal-in-time dynamic system proposed for the effective modeling of complex diffusive regimes in heterogeneous media. We present its solutions and their commonly studied statistics such as the mean square distance. This interesting model employs a nonlocal operator to replace the conventional first-order time-derivative. It introduces a finite memory effect of a constant length encoded through a kernel function. The nonlocal-in-time operator is related to fractional time derivatives that rely on the entire time-history on one hand, while reduces to, on the other hand, the classical time derivative if the length of the memory window diminishes. This allows us to demonstrate the effectiveness of the nonlocal-in-time model in capturing the crossover widely observed in nature between the initial sub-diffusion and the long time normal diffusion.

Key words. Nonlocal model, nonlocal operators, mean square displacement, sub-diffusion, numerical methods.

1. Introduction

Diffusion is one of the prominent transport mechanisms in nature. A conventional normal diffusion typically refers to a diffusion process with the mean squared displacement changing linearly in time. Over the last few decades, anomalous diffusion has attracted a lot of attention due to its associations with many diffusion processes in heterogeneous media [4, 28], though its origins and relevant mathematical models can take significantly different forms [23, 33, 3, 26]. On one hand, new experimental standards have been called for [32] to obtain more fundamental statistics on anomalous diffusion processes. On the other hand, there are needs for in-depth studies of novel models, which might be non-conventional and nonlocal [33, 10].

The goal of this work is to present a simple dynamic equation that provides an effective description of the diffusion process encompassing different diffusive regimes. This is motivated by some recent experimental reports on the crossover between initial transient sub-diffusion and long time normal diffusion in various settings [17, 36]. The main feature of the proposed model is to incorporate memory effect or time correlations with a finite and fixed horizon length, denoted by $\delta > 0$, across the dynamic process. The memory kernel is constant in space and time so that neither spatial inhomogeneities nor time variations get introduced in the diffusivity coefficients. This is different from the approaches taken in other models of anomalous diffusion like the variable-order fractional differential equation [35, 40, 38] or the diffusing diffusivity model [6, 16, 18]. Very recently, [2] presented an interesting study on the use of different types of fractional differential equations to capture crossovers from the initial superdiffusive regime to later normal and subdiffusive regime. Our findings presented here complement those given in [2] as we are aiming to study the opposite process of changing from subdiffusive to normal regimes. On one hand, our nonlocal in time model is quite elegant: it does

Received by the editors on September 26, 2022 and, accepted on January 6, 2023.
2000 *Mathematics Subject Classification.* 35R35, 49J40, 60G40.

not require the introduction of a time dependent memory kernel, and it does not involve spatial heterogeneities. On the other hand, the model under consideration here is also very intuitive: due to the fixed memory span, the memory effect plays an dominant role in the whole or a large part of the time history so that the non-Markovian effect is evident during the early time period. As time goes on, however, the fixed memory span becomes less and less significant in comparison with the longer and longer life history. Hence, the process becomes more and more Markovian like. As a result, the transition from sub-diffusion to normal diffusion occurs naturally. A rigorous demonstration of this intuitive picture will be given later in more details. We note that while the nonlocal-in-time diffusion equation may be related to fractional diffusion equations [29, 30, 33] by taking special memory kernels [1], they in general provide a new class of models that effectively serve as a bridge between anomalous diffusion and normal diffusion, with the latter being a limiting case as the horizon length $\delta \to 0$.

Specifically, let Δ be the Laplacian (diffusion) operator in the spatial variable x and $g = g(x,t)$ represent the initial (historical) data, we consider the following nonlocal-in-time diffusion equation for $u = u(x,t)$:

$$\begin{aligned} \mathcal{G}_\delta u &= \Delta u \qquad \forall\ t > 0, \\ u &= g \qquad\ \ \forall\ t \in (-\delta, 0). \end{aligned} \tag{1}$$

The nonlocal operator $\mathcal{G}_\delta$ in (1) is defined by

$$\mathcal{G}_\delta v(t) = \int_0^\delta \frac{v(t) - v(t-s)}{s} \rho_\delta(s)\, ds, \tag{2}$$

where the memory kernel function $\rho_\delta = \rho_\delta(s)$ is assumed to be nonnegative with a compact support in $(0,\delta)$ and is integrable in $(0,\delta)$. In case that $s^{-1}\rho_\delta(s)$ is unbounded only at the origin, the integral in (2) should be interpreted as the limit of the integral of the same integrand over (ϵ,δ) for $\epsilon > 0$ as $\epsilon \to 0$, where such a limit exists in an appropriate mathematical sense.

The nonlocal operator $\mathcal{G}_\delta$ has been discussed in [13, 11, 12], and it forms part of the nonlocal vector calculus [9, 27]. The positive nonlocal horizon parameter δ appearing in (2) represents the memory span. For suitably chosen kernels, as $\delta \to 0$, nonlocal and memory effects diminish, so that the zero-horizon limit of the nonlocal operator $\mathcal{G}_\delta u$ corresponds to the standard first order derivative $\frac{d}{dt}u$. In particular, under the normalization condition

$$\int_0^\delta \rho_\delta(s)\, ds = 1, \tag{3}$$

the kernel function $\rho_\delta(s)$ can be seen as a probabilistic density function(PDF) defined for s in $(0,\delta)$, so that the nonlocal operator (2) can be viewed as a "continuum" average of the backward difference operators over the memory span measured by the horizon $\delta > 0$. The local limit, i.e., the standard derivative, is simply the extreme case where $\rho_\delta(s)$ degenerates into a singular point measure at $s = 0$. In fact, one can see from a formal Taylor expansion that

$$\mathcal{G}_\delta v(t) = \frac{dv}{dt}(t) + \sum_{k=2}^{\infty} \frac{v^{(k)}(t)}{k!} \int_0^\delta (-s)^{k-1} \rho_\delta(s)\, ds = \frac{dv}{dt}(t) + O(\delta) \to \frac{dv}{dt}(t).$$

as $\delta \to 0$. Then the nonlocal-in-time model (1) recovers the following classical (local) diffusion model

$$\partial_t u = \Delta u \qquad \forall\ t > 0, \quad \text{with } u(0) = g(0).$$

Meanwhile, in another extreme case that $\delta \to \infty$, we may let the initial (historical) data $g(x,t) := g(x)$ for all $t \in (-\infty, 0)$, and take the kernel function to be of the fractional type, i.e.,

$$\rho_\delta(s) = \frac{\alpha}{\Gamma(1-\alpha)} s^{-\alpha} \chi_{(0,\delta)}(s), \quad \text{for some } \alpha \in (0,1), \tag{4}$$

where $\chi = \chi_{(0,\delta)}(s)$ denotes the characteristic (indicator) function of $(0,\delta)$. Then, as $\delta \to \infty$, the nonlocal operator $\mathcal{G}_\delta$ reproduces the Marchaud fractional derivative of order α, which is equivalent to the Caputo fractional derivative for smooth functions at time $t > 0$ (e.g. [22, p. 91]). As a consequence, the nonlocal-in-time diffusion model (1) recovers the fractional sub-diffusion model for $u = u(x,t)$:

$$\partial_t^\alpha u - \Delta u = 0, \quad \forall\, t > 0, \quad \text{with } u(0) = g.$$

The fractional subdiffusion model has often been used to describe the continuous time random walk (CTRW) of particles in heterogeneous media, where trapping events occur randomly. In particular, particles get repeatedly immobilized in the environment for a trapping time drawn from the waiting time probability density function with a heavy tail, i.e., $\omega(t) \propto t^{-\alpha-1}$ as $t \to \infty$. [29, 30]. As discussed later in Section 3, the nonlocal-in-time model (1) can be also related to a trapping model where the kernel function ρ_δ describes the distribution of waiting time probability.

In comparison with the classical diffusion equation (the $\delta \to 0$ local limit) and fractional diffusion (the $\delta \to \infty$ limit), the nonlocal-in-time evolution equation provides a more general model and an interesting intermediate case to study the "finite history dependence" with a given $\delta \in (0, \infty)$. The goal of this paper is to present the modeling capability and the behavior of the solutions of the nonlocal-in-time dynamics (1), so as to demonstrate how the simple PDE model can serve as a bridge linking the two limiting scenarios described by the standard normal diffusion and the fractional sub-diffusion respectively. We refer to [13, 11, 27, 12] for more development on the mathematical background and numerical analysis of the nonlocal operators and the nonlocal-in-time dynamic systems. We note that some of the materials presented in this work appeared in [8], which was based on an unpublished version of the results presented here, as acknowledged explicitly in [8]. The materials are used here with permission.

2. Solutions of the nonlocal-in-time model

To begin with, we shall study the solution behaviour of the nonlocal-in-time diffusion dynamics. Throughout the section, we choose fractional type memory kernels of the type (4) as illustrations. This also allows us to make comparisons with the local and fractional diffusion models. To correlate with data observed from physical and biological experiments, we focus on the spatial solution profiles at various stages and the time evolution of the mean square displacement.

2.1. Fundamental solutions. Our starting point the model (1) defined spatially on the real line, with a time nonlocal initial distribution given by a time stationary Dirac-delta measure in x at the origin, denoted by $\delta(x)$. That is, we set in (1) that $g(x,t) = \delta(x)$ for $x \in (-\infty, \infty)$ and $t \in (-\delta, 0)$. Under this setting, the solution to (1) is usually called the fundamental solution, following the convention of local in time PDE models. In particular, we consider kernel functions that have a unit integral, as assumed in (3). Then, as $\delta \to 0$, the limit of the solution indeed gives the well-known fundamental solution of the heat equation, expressed by a Gaussian

function

$$u_0(x,t) = \frac{1}{\sqrt{4\pi t}} e^{-x^2/4t}.$$

As a comparison, in another limit case where δ approaches ∞, with the kernel function (4), the nonlocal-in-time operator leads to a fractional derivative. In other words, $\mathcal{G}_\delta u(t) \to \mathcal{G}_\infty u(t) = \partial_t^\alpha u(t)$. In this limit, the fundamental solution becomes the Fox H function

$$u_\infty(x,t) = \frac{1}{2} t^{-\alpha/2} P_\alpha(x/t^{\alpha/2}),$$

with $P_\alpha(y) \sim A y^a e^{-by^c}$ as $y \to \infty$ for constants $a = \frac{2\alpha-2}{2-\alpha}$, $b = (2-\alpha)2^{-\frac{2}{2-\alpha}}\alpha^{\frac{\alpha}{2-\alpha}}$, and $c = \frac{2}{2-\alpha}$ [25].

For the nonlocal-in-time diffusion model (1), let $\widetilde{u}(\xi,t)$ denote the Fourier transform of fundamental solution $u(x,t)$ with respect to x, we get a scalar nonlocal-in-time initial value problem

$$\mathcal{G}_\delta \widetilde{u}(\xi,t) + \xi^2 \widetilde{u}(\xi,t) = 0$$

for $t > 0$ with an initial data $\widetilde{u}(\xi,t) = 1$ for $t \in (-\delta, 0)$. The Laplace transform of $\widetilde{u}(\xi,t)$ with respect to time t is given by

$$\widehat{\widetilde{u}}(\xi,z) = \frac{z^{-1}K(z)}{K(z)+\xi^2}, \text{ with } K(z) = \int_0^\delta (1-e^{-zs}) s\rho_\delta(s)\, ds.$$

Applying the inverse Laplace and Fourier transforms, we obtain a formal analytic representation of the solution

$$\begin{aligned} u(x,t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathrm{i}\xi x} \frac{1}{2\pi\mathrm{i}} \int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} \frac{z^{-1}K(z)}{K(z)+\xi^2} e^{zt}\, dz\, d\xi \\ &= \frac{1}{2\pi\mathrm{i}} \int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} e^{zt} \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathrm{i}\xi x} \frac{z^{-1}K(z)}{K(z)+\xi^2}\, d\xi\, dz \\ &= \frac{1}{2\pi\mathrm{i}} \int_{\sigma-\mathrm{i}\infty}^{\sigma+\mathrm{i}\infty} e^{zt} \widehat{u}(x,z)\, dz. \end{aligned}$$

The above integrals are well-defined for $(x,t) \in \mathbb{R} \times (0,\infty)$. Moreover, we can observe that the fundamental solution $u = u(x,t)$ is continuous in x and t and piecewise smooth in x away from $x = 0$ for $t > 0$. For $z > 0$, we can get

$$\widehat{u}(x,z) = \frac{\sqrt{K(z)}}{2z} e^{-|x|\sqrt{K(z)}}. \tag{5}$$

Concerning the time variation of the fundamental solution of (1) with the kernel (7), for example at $x = 0$, we note first $\widehat{u}(0,z)$ is monotone decreasing in z and

$$\widehat{u}(0,z) \approx \begin{cases} \sqrt{\frac{\Gamma(2-\alpha)}{4\alpha\delta^{1-\alpha}}} z^{\alpha/2-1} & \text{as } z \to \infty, \\ \frac{1}{2} z^{-1/2} & \text{as } z \to 0. \end{cases}$$

Then the Karamata-Feller-Tauberian theorem [14] gives

$$u(0,t) \approx \begin{cases} \frac{\sqrt{\Gamma(2-\alpha)}}{\sqrt{4\alpha\delta^{1-\alpha}}\Gamma(1-\alpha/2)} t^{-\alpha/2} & \text{as } t \to 0, \\ \frac{1}{\sqrt{4\pi t}} & \text{as } t \to \infty. \end{cases}$$

Again, as $t \to \infty$, the decay property remains to be the same as the standard heat kernel, while as $t \to 0$, the solution behaves like $O(t^{-\alpha/2})$, which matches with the behavior of the fundamental solution for the fractional sub-diffusion. This observation is already very telling about the model (1) in capturing the crossover of diffusion regimes.

As a more vivid illustration, Fig 1 shows a plot of the time evolution of the numerical solution $u(x,t)$ evaluated at $x=0$ for $\alpha=0.2$ and $\delta=0.1$. The tangents with different slopes, corresponding to $-\alpha/2=-0.1$ and $-1/2$ are also plotted Fig 1 to further illustrate the two limiting regimes.

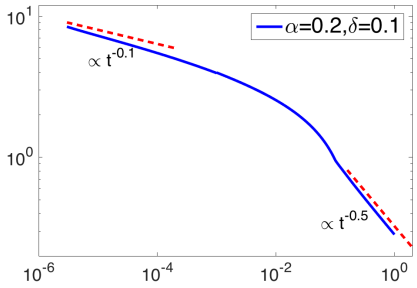


FIGURE 1. $u(0,t)$ of nonlocal-in-time diffusion with $\delta=0.1$, $\alpha=0.2$.

2.2. Mean square displacement (MSD). One of the most common characteristics of the diffusion process is the mean square displacement (MSD), denoted by $m(t)$ and defined by

$$m(t) := \int_{\mathbb{R}} x^2 u(x,t)\,dx.$$

As in the previous discussion, we assume that the kernel ρ_δ is a normalized probability density function (PDF), i.e., (3) is satisfied.

By the expression of the fundamental solution (5), it is easy to see the corresponding MSD $m(t)$ satisfies satisfies the nonlocal initial value problem

$$\mathcal{G}_\delta m(t) = 2 \;\forall\; t>0, \quad \text{with } m(t)=0 \;\forall\; t\in[-\delta,0], \tag{6}$$

One can get that $m(t)\approx 2t$ as $t\to\infty$. Meanwhile, if we select special kernels of the fractional type:

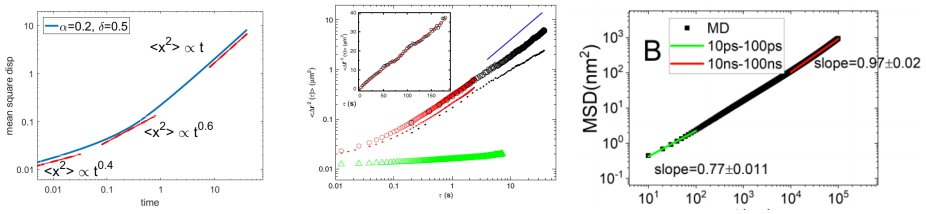
$$\rho_\delta(s) = (1-\alpha)\delta^{\alpha-1}s^{-\alpha}, \tag{7}$$

then we have the asymptotic behaviour

$$m(t) \approx \frac{2\sin(\alpha\pi)\delta^{1-\alpha}}{(1-\alpha)\pi}t^\alpha \quad \text{as } t\to 0.$$

Note that when $\alpha\to 1$ the corresponding MSD $m(t)$ becomes $2t$. We again observe the transition from the sub-diffusion initially to normal diffusion at a later time.

In figure 2, we plot the numerical solution of (6), i.e., the mean square displacement of the nonlocal model corresponding to $\rho_\delta(s)$ given by (7) with $\alpha=0.2$ and $\delta=0.5$. The initial power-law scaling with exponent $2\alpha=0.4$ and the later linear scaling, as well as the transition in between are also shown to better illustrate the changes in different regimes. The simulated experiment further illustrates the analytically suggested change from the early fractional anomalous diffusion regime to the later normal diffusion regime. This type of "transition" or "crossover" behavior have been captured by a number of experimental observations, e.g. diffusion in lipid bilayer systems of varying chemical compositions [19, Figure 2], hydration water molecules [36, Figure 4B] and lateral motion of the acetylcholine receptors on live muscle cell membranes[17, Figs. 3, 4] , as shown in the reproduced plots in figure 2.



(a) (b) (c)

FIGURE 2. (a) The plot of MSD of nonlocal-in-time diffusion model (1) with $\alpha = 0.2$ and $\delta = 0.5$ (A similar figure with a different time range has also been presented in [8, Section 6.3]); (b) The immobile AChRs: Crossover from sub-diffusion to normal diffusion observed from the MSD curve [17, Figure 3] reproduced here with permission. (c) MSD derived from MD (black solid square) simulations of hydration water molecules. Green and red lines represent power-law fits in time window from 10ps to 100ps and from 10ns to 100ns respectively [36, Figure 4B] reproduced here with permission.

Based on the earlier discussion, we know that if the initial condition g is history-independent, i.e., $g(x,t) := u_0(x)$, then $m(t)$, the mean square displacement of the nonlocal-in-time diffusion with a kernel function (7), is increasing and exhibits a weak singularity in the beginning, the same as its fractional counterpart. We can further study the dependence of MSD on the initial historical data by changing the history-dependent initial data to the nonlocal-in-time model (1). In Figure 3, we plot the solution of initial value problem $\mathcal{G}_\delta m(t) = 2$, with different initial (historical) data for $t \in (-\delta, 0)$:

$$g_1(t) = 5(1+2t), \quad g_2(t) = \frac{1+2t}{2} \quad \text{and} \quad g_3(t) = 10 * \chi_{[-\delta,\delta/2]}.$$

Our numerical results show that the solution $m(t)$ for the linear initial data g_1 decays in the beginning, due to the historical-dependence of the nonlocal-in-time dynamics (where the initial data grows rapidly), and then increases at a later time. For the initial data g_2, which is also linear in time but with a smaller slope than the previous case, we observe a strictly increasing mean square displacement function. In fact, one can prove that for initial data $k(1+2t)$, the solution $m(t)$ will keep strictly increasing when $k \leq 1$. Meanwhile, for the step-like initial data g_3, the solution increases dramatically near $t = 0$, and then decreases slightly before its constantly linear growth. Through this simple experiment, we see that the growth behavior of $m(t)$ for various historical initial data can show quite different characteristics. This merits further theoretical investigation.

2.3. More on the long time normal diffusion limit. The long-time normal diffusion behavior can also be observed directly from the model (1) by a simple scaling. In particular, for a typical rescaled density $\rho_\delta(s) = \delta^{-1}\rho_1(\delta^{-1}s)$ with ρ_1 being a density over the unit interval, then for δ_0 and $v(x,t) = u(x'.t')$ with

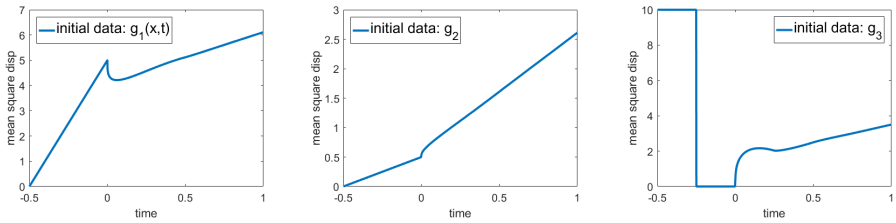


FIGURE 3. Mean square displacement of nonlocal-in-time model with different initial distribution, $\alpha = 0.2$, $\delta = 0.5$.

$x' = x\delta_0^{-1/2}$ and $t' = t\delta_0^{-1}$, it holds that

$$\begin{aligned}
\mathcal{G}_\delta v(x,t) &= \int_0^\delta \frac{v(x,t) - v(x,t-s)}{s}\rho_\delta(s)\,ds \\
&= \frac{1}{\delta}\int_0^\delta \frac{v(x,t) - v(x,t-s)}{s}\rho_1(\frac{s}{\delta})\,ds \\
&= \frac{1}{\delta}\int_0^\delta \frac{u(\frac{x}{\sqrt{\delta_0}},\frac{t}{\delta_0}) - u(\frac{x}{\sqrt{\delta_0}},\frac{t-s}{\delta_0})}{s}\rho_1(\frac{s}{\delta})\,ds \\
&= \frac{1}{\delta_0}\int_0^{\delta'} \frac{u(x',t') - u(x',t'-s')}{s'}\rho_{\delta'}(s')\,ds' \\
&= \frac{1}{\delta_0}\mathcal{G}_{\delta'}u(x',t'),
\end{aligned}$$

by a change of variables $\delta = \delta'\delta_0$ and $s = s'\delta_0$. Then, using $\Delta_x v(x,t) = \delta_0^{-1}\Delta_{x'}v(x',t')$, we see that for any solution $u = u(x,t)$ of the nonlocal in time diffusion model (1), the rescaled solution v satisfies the same equation in the rescaled variables corresponding to a kernel with a rescaled horizon $\delta' = \delta/\delta_0$ (and rescaled initial data).

It is then interesting to consider some limiting cases. For example, if we let $\delta_0 \to \infty$, then the rescaled horizon δ' goes to zero. In this case, we know that $\mathcal{G}_{\delta'}$ is effectively the conventional local time derivative. We thus see that with a diminishing memory effect, v approximately satisfies the classical normal diffusion on the $O(t)$ scale, which also implies that u effectively behaves like the normal diffusion on the long time scale $O(t') >> O(t)$.

For a fixed nonlocal horizon δ and at a given time t, one may derive the asymptotic behavior of the solution to the nonlocal-in-time model (1), as $x \to \infty$, from its Laplace transform. In particular, we have

$$\widehat{u}(x,z) \approx \frac{1}{2}c_\delta z^{\alpha/2-1}e^{-c_\delta|x|z^{\alpha/2}}, \quad \text{as} \quad z \to \infty,$$

for $c_\delta = \sqrt{\Gamma(2-\alpha)\delta^{\alpha-1}\alpha^{-1}}$. Hence by inverse Laplace transform, we have

$$u(x,t) \approx \frac{1}{2}c_\delta t^{-\alpha/2}P_{\alpha/2}(c_\delta x/t^{\alpha/2}), \quad \text{as} \quad t \to 0,$$

with the Fox H function P_α. This implies that $u(x,t)$ has a fractional exponential tail, i.e., for a fixed $t > 0$,

$$u(x,t) \approx A_{\alpha,\delta,t}e^{-b_{\alpha,\delta,t}x^{2/(2-\alpha)}}, \quad \text{as} \quad x \to \infty.$$

2.4. Smoothing properties. In addition to statistics like the MSD, another interesting feature of the nonlocal-in-time dynamics is its gradual smoothing property. The latter refers to the fact that the solutions can become more and more smooth

(as functions of the spatial variables) as time goes on. Such a property can be derived similarly as done in [13]. In fact, with the fractional kernel function in (7), we may see that

$$|\widetilde{u}(\xi,t)| \le c/(1+b_\delta|\xi|^2 t^\alpha)$$

for any $t>0$ and some constants c and b_δ [13, Theorem 3.2]. This implies that $|\xi|^s|\widetilde{u}(\xi,t)|$ is square-integrable in $(-\infty,\infty)$ only if $s<\frac{3}{2}$. This restriction reflects the limited smoothness of the fundamental solution, i.e., the solution $u(x,t)$, for any $t>0$, has nearly 3/2-order square integrable fractional spatial derivative and hence is only piecewise smooth. In fact, it roughly gains two more orders of differentiability with each additional δ increment in time. This observation indicates that the smoothing of the solution $u(x,t)$ takes place incrementally over time.

2.5. Comparison with fractional and normal diffusion via numerical illustrations. In Figure 4, we plot the numerical solution of the nonlocal-in-time model with $\delta=0.1$, $\alpha=0.3$, and compare with solutions of the related fractional diffusion and normal diffusion. The initial data is the Dirac-delta measure concentrated at the origin. We use the numerical solution in the finite domain $(-50,50)$ with Dirichlet boundary condition to approximate the numerical solution in the infinite domain $(-\infty,\infty)$.

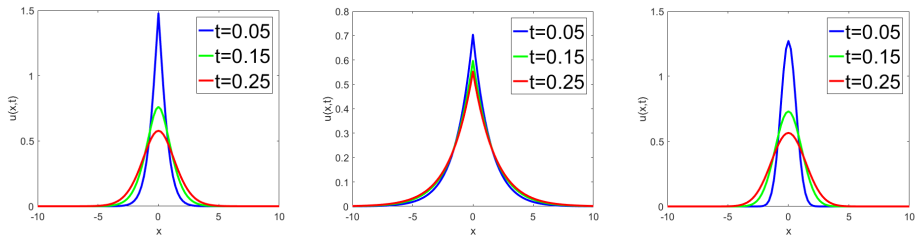
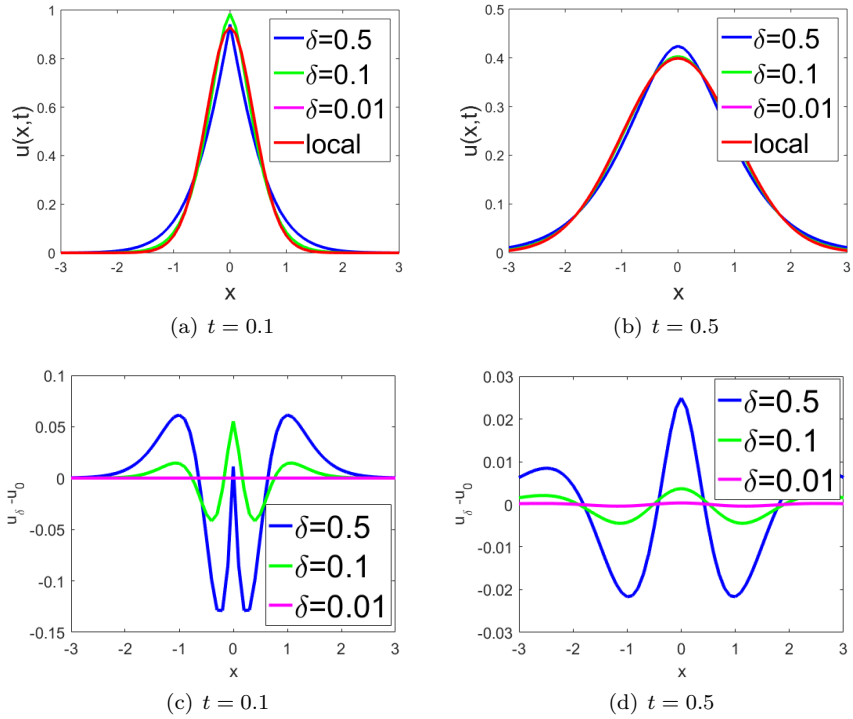


FIGURE 4. Left: Numerical solutions of the nonlocal-in-time diffusion equation with $\delta=0.1$, $\alpha=0.3$. Middle: Numerical solutions of the fractional diffusion with $\alpha=0.3$. Right: Numerical solutions of the local diffusion.

From numerical experiments, we observe that the solution of the nonlocal in time model is only spatially piecewise smooth and is getting more regular as t increases (the green and red curves in Figure 4 (left)). This interesting observation shows that, as another distinct feature of the finite memory effect of the nonlocal-in-time model, the solution has the same smoothing property as the corresponding fractional sub-diffusion (Figure 4 (middle)) initially, but gets improved smoothing property at a later time, and finally exhibits the similar smoothing behaviour of normal diffusion (Figure 4 (right)) as t tends to infinity.

In above comparison, the kernel in (7) differs from the ρ_∞ used for the fractional diffusion up to a constant factor. If a new kernel is taken as $\nu\rho_\delta$ with a constant factor $\nu>0$, then the fundamental solution of the nonlocal model with the new kernel is $u(x/\sqrt{\nu},t)$, with $u(x,t)$ being the fundamental solution of (1) with the original kernel.

It is interesting to numerically test the local limit of the nonlocal-in-time model. In Figure 5, we plot the numerical solutions of nonlocal-in-time model with different nonlocal horizons at $t=0.1$ and 0.5, as well as the difference between the nonlocal



(a) $t = 0.1$ (b) $t = 0.5$

(c) $t = 0.1$ (d) $t = 0.5$

FIGURE 5. Figures (a) and (b): The plot of numerical solutions of nonlocal-in-time diffusion at $t = 0.1$, 0.5 with $\alpha = 0.5$ and different nonlocal horizons. Figures (c) and (d): The plot of difference between solutions of nonlocal diffusion with different nonlocal horizons and the solution of local diffusion.

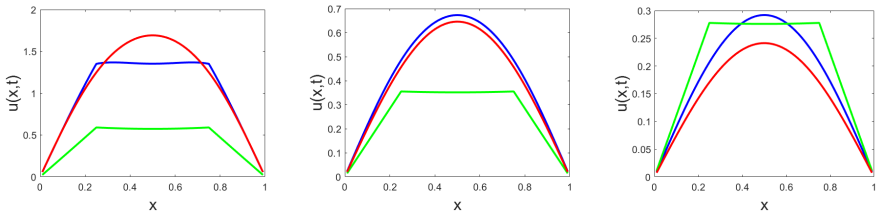
solutions and the local one. It can be observed that as δ goes to zero, the solution of the nonlocal diffusion model converges to the solution of the local one.

2.6. Nonlocal-in-time dynamics on a finite spatial domain. To complement the study on the infinite spatial domain, we now examine the nonlocal-in-time parabolic equation (1) in a bounded domain. Specifically, we consider the following initial-boundary value problem for $u = u(x,t)$:

$$\begin{aligned} \mathcal{G}_\delta u(x,t) - \Delta u(x,t) &= 0, && \text{in } \Omega \times [0,T], \\ u(x,t) &= 0, && \text{in } \partial\Omega \times [0,T], \\ u(x,t) &= g(x,t), && \text{in } \Omega \times [-\delta, 0), \end{aligned} \tag{8}$$

where Ω is a bounded convex polyhedral domain in $\mathbb{R}^d$ ($d \geq 1$) with a boundary $\partial\Omega$. Theoretical studies of the initial-boundary value problem (8), such as well-posedness, regularity estimate, numerical analysis, and asymptotic behaviour, can be found in [13].

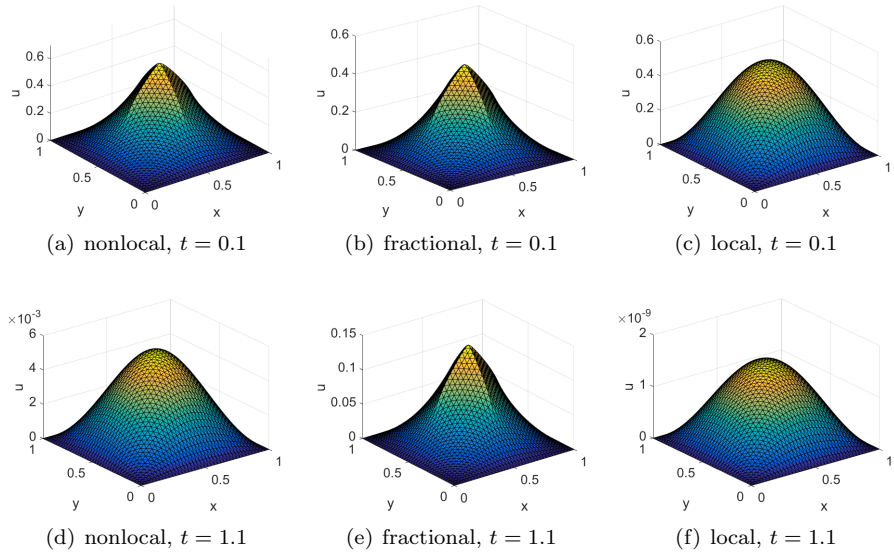
To illustrate our findings, we use the fractional kernel given by (7) and compare its solution behavior with those of fractional diffusion and local diffusion. In Figure 6, we present a numerical solution of a 1-D nonlocal model (8) in the unit interval $\Omega = (0,1)$, at different time, with $\delta = 0.1$ and $\alpha = 0.5$. In our computation, the initial (historical) data is taken as a Dirac-delta measure concentrated at



(a) $t = 0.05$ (b) $t = 0.15$ (c) $t = 0.25$

FIGURE 6. Numerical solutions of local (red curve), fractional (green curve) and nonlocal (blue curve) diffusion models, $\delta = 0.1$ and $\alpha = 0.5$.

$x = 1/4$ and $x = 3/4$. We observe that the nonlocal diffusion gradually regularize the solution as t increases. Moreover, the fractional diffusion decays fastest for small t, and the classical local diffusion decays fastest for large t, while nonlocal diffusion exhibits an intermediate behavior.



(a) nonlocal, $t = 0.1$ (b) fractional, $t = 0.1$ (c) local, $t = 0.1$

(d) nonlocal, $t = 1.1$ (e) fractional, $t = 1.1$ (f) local, $t = 1.1$

FIGURE 7. Numerical solutions of nonlocal (left), fractional (middle), local (right) models with $\Omega = (0,1)^2$, $\delta = 1$ and $\alpha = 0.5$ at different time.

Finally, we show a two-dimensional example with $\Omega = (0,1)^2$ and an initial data given by the Dirac-delta measure concentrated on the cross $(0.25, 0.75) \times \{0.5\} \cup \{0.5\} \times (0.25, 0.75)$. Comparisons with the corresponding solutions of the fractional and local diffusion models are all provided.

3. Additional discussions on nonlocal-in-time models

As we can see from the examples presented in this paper, the nonlocal-in-time model (1) is a simple modification of the traditional dynamic systems where the instant rate of change ($\partial_t u$) is replaced with a nonlocal rate of change ($\mathcal{G}_\delta u$). The

local limit, i.e., the standard derivative, is simply the extreme case where the kernel function $\rho_\delta(s)$ in (2) degenerates into a singular point measure at $s = 0$. On the other hand, we have known that by picking suitable fractional type kernels with $\delta \to \infty$, the nonlocal-in-time model can recover fractional subdiffusion equations as well. We thus advocate the equation (1) as a more general model which build an bridge between classical diffusion and time-fractional subdiffusion equations.

There have been a large number of studies on time-fractional dynamics, which are used to model the continuous time random walk of particles in heterogeneous media [29, 4, 30]. The fractional derivative arising in fractional sub-diffusion model results from a fractional power-law type waiting time probability. However, such specific choices of memory kernels are rather artificially restrictive. One may argue that perhaps there remains a lack of compelling evidence that the nature is confined by such limited forms of kernels.

For example, let us consider the situation where particles may explore some environment that is a kind of labyrinthine (bearing short-time sub-diffusion) but homogenizes on large length scales (resulting in a long-time normal diffusion). In recent experimental studies of such systems, there have been a number reports on different diffusion regimes from short-time sub-diffusion to long-time normal diffusion, e.g., [19, 17]. This motivates people to look for a simple and effective model to capture the transient dynamics.

Naturally, there might be different mathematical models which are able to capure such crossover phenomena. As an example, besides the recently proposed diffusing diffusivity model (cf. [6, 16, 18]), a variable-order fractional diffusion equation

$$\partial_t^{\alpha(t)} u - \Delta u = f \tag{9}$$

has been used in [35] with a time-dependent order $\alpha(t)$. However, analyzing such model (theoretically or numerically) remains a difficult task from a mathematical perspective, besides the challenge for further physical validation. See also [40, 38] for an alternative variable-order subdiffusion model.

Another popular model to capture the crossover between subdiffusion and normal diffusion is the time-tempered fractional diffusion equation [39, 15, 31]

$$\partial_t^{\alpha,\sigma} u - \Delta u = f \tag{10}$$

where the time-tempered fractional derivative is defined as $\partial_t^{\alpha,\sigma} u(t) := e^{-\sigma t}\partial_t^\alpha [e^{\sigma t} u(t)]$ with constants $\sigma > 0$ and $\alpha \in (0, 1)$. Note that the historical effect of the model (10) dates back to $t = 0$ and its smoothing effect is always second-order in Sobolev spaces for all $t > 0$.

In contrast, the nonlocal-in-time model (1) with a constant finite history dependence leads to a more direct and intuitive way to interpret and to capture the crossover behavior. Moreover, although different models may produce similar behavior on MSD, other statistics and spatial/temporal patterns of the solutions may differ [28, 34]. Thus, it is important to conduct more mathematical investigations and numerical simulations, like the studies presented here, in order to provide a deeper understanding of the underlying process.

From a modeling perspective, the nonlocal-in-time model (1) can also be explained by a model of random walk with trapping events driven by a given probability intensity. For a truncated fractional kernel function of the type (7), the probability intensity of the trapping event is related to $c_\alpha s^{-\alpha-1}\chi_{(0,\delta)} =: \rho_\delta(s)/s$. As an illustration, let us focus on a simple 1-D case as follows.

In discrete time steps with a uniform step size τ, a random walker is assumed to stay in its current position, or move to one of its nearest neighbor sites with length

h and in a random left or right direction. Let $u(x,t)$ be the probability that the walker appears at position x and time t, $\eta(x,t)$ be the probability that the walker has just arrived at x and t, and ω_k is the probability that the walker waits for at least $k\tau$ after arriving at a position.

Under the assumption that the maximal waiting time is δ, we then derive

$$u(x,t) = \sum_{k=1}^{\delta/\tau} \omega_k \eta(x, t-k\tau) + \eta(x,t), \tag{11}$$

We let p_k be the probability that the walker stops exactly $k\tau$ after its arrival, which for $k = 1, 2, ..., \delta/\tau$, is assumed to be of the form:

$$p_{k-1} = \Big(\frac{1}{k\tau}\int_{(k-1)\tau}^{k\tau} s^{-\alpha}\, ds\Big) \Big/ \Big(\sum_{k=1}^{\delta/\tau} \frac{1}{k\tau}\int_{(k-1)\tau}^{k\tau} s^{-\alpha}\, ds\Big).$$

Then, it is obvious that

$$\omega_k = 1 - \sum_{i=0}^{k-1} p_k.$$

The discrete model of random walk with trapping events gives that for h small,

$$\begin{aligned} \eta(x,t) &= \textstyle\sum_{k=1}^{\delta/\tau} \frac{1}{2}(\eta(x-h, t-k\tau) + \eta(x+h, t-k\tau))p_{k-1} \\ &\approx \textstyle\sum_{k=1}^{\delta/\tau} (\eta(x, t-k\tau) + \frac{h^2}{2}\eta_{xx}(x, t-k\tau))p_{k-1}. \end{aligned}$$

Let $\mathcal{G}_\delta^\tau$ be a discrete nonlocal-in-time operator defined by

$$\mathcal{G}_\delta^\tau u(t) = u(t)\Big(\textstyle\sum_{k=1}^{\delta/\tau} \frac{1}{k\tau}\int_{(k-1)\tau}^{k\tau} \rho_\delta(s)\, ds\Big) - \sum_{k=1}^{\delta/\tau} \frac{u(t-k\tau)}{k\tau}\int_{(k-1)\tau}^{k\tau} \rho_\delta(s)\, ds, \tag{12}$$

with $\rho_\delta(s) = (1-\alpha)\delta^{\alpha-1}s^{-\alpha}$. Then a simple calculation by Taylor's expansion yields that $\mathcal{G}_\delta^\tau u(t) \approx \mathcal{G}_\delta u(t)$ for small τ, and for

$$c_{\delta,\alpha} = (1-\alpha)\delta^{\alpha-1}\sum_{k=1}^{\infty}\frac{1}{k}\int_{k-1}^{k} s^{-\alpha}\, ds$$

we can derive that

$$\sum_{k=1}^{\delta/\tau}\frac{1}{k\tau}\int_{(k-1)\tau}^{k\tau}\rho_\delta(s)\, ds \approx c_{\delta,\alpha}\tau^{-\alpha}.$$

Therefore, we observe that for for small $\tau^\alpha \approx h^2$,

$$\mathcal{G}_\delta^\tau \eta(x,t) - \frac{c_{\delta,\alpha}}{2}\frac{h^2}{\tau^\alpha}\eta_{xx}(x,t) = 0. \tag{13}$$

The nonlocal-in-time model (1) also follows from (13) by applying $\mathcal{G}_\delta^\tau$ on the equation (11), using (13) and letting $\tau^\alpha \approx h^2 \to 0$.

To substantiate this stochastic explanation, we compared numerical solution of (13) computed by finite different method with the Monte-Carlo solution using the particle method, in Figure 8. We consider the infinite domain in one space dimension $\Omega = \mathbb{R}$, and let the initial (historical) data be the Dirac-delta measure concentrated at $x = 0$ i.e., we only track the movement of a single particle, starting from the position $x = 0$. Numerical results indicate that the Monte-Carlo solution fits the finite difference solution very well, and this supports our theoretical results.

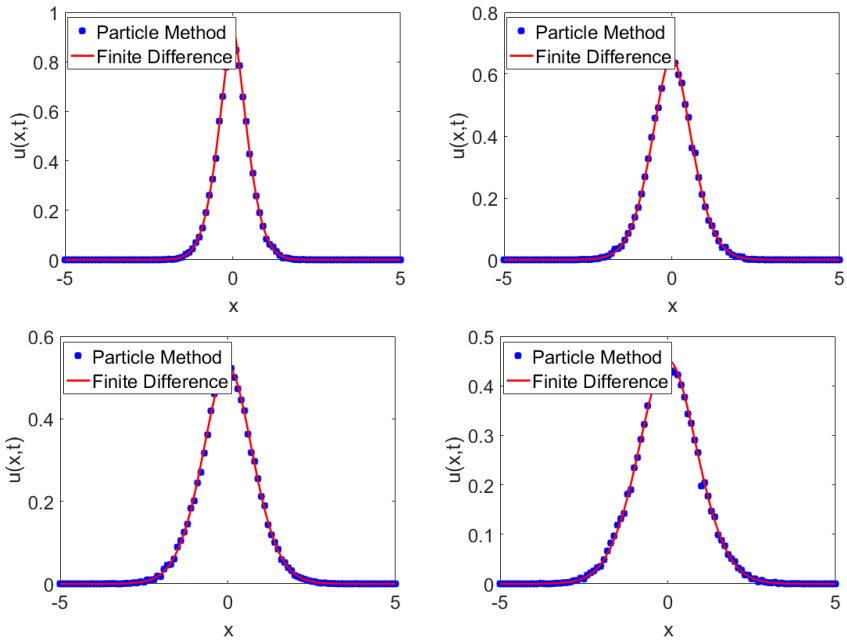


FIGURE 8. Numerical solutions of the nonlocal-in-time diffusion equation with $\delta = 0.2$, $\alpha = 0.75$ and Dirac-Delta initial condition, at $t = 0.1$, 0.2, 0.3 and 0.4, respectively.

4. Numerical approximation and challenges

Sine the closed-form solution of the nonlocal-in-time model (1) is not available in general, it is important to design efficient numerical methods to approximately solve it. In this section, we shall briefly introduce some numerical methods for initial-boundary value problem (8), with the fractional-type kernel function defined as (7), and discuss the challenge in computation and analysis of the resulting discrete system. See the monographs [37] and [21] for the numerical treatment for its normal diffusion and fractional subdiffusion counterparts, respectively.

As we introduced in Section 2, one distinct feature of the nonlocal-in-time diffusion is the gradual smoothing effect in space, which only allow low-order spatial approximation for small t due to the limited regularity, in case of nonsmooth initial condition, but might allow high-order spatial approximation for large t. Following the standard setting in [37, 20], we apply the finite element method (FEM) in space. Let $\mathcal{T}_h$ be a shape regular quasi-uniform triangulation of the domain Ω into d-simplexes, with a mesh size h. Over $\mathcal{T}_h$, let X_h be the finite element space consisting of continuous piecewise r-th order polynomials. Then the semidiscrete Galerkin FEM for problem (8) reads: find $u_h(t) \in X_h$ such that

$$\begin{aligned} (\mathcal{G}_\delta u_h, \chi) + (\nabla u_h, \nabla \chi) &= (f, \chi), \qquad \forall\, \chi \in X_h,\ T \geq t > 0, \\ u_h(t) &= P_h g, \end{aligned} \tag{14}$$

where $P_h : L^2(\Omega) \to X_h$ denotes the L^2-projection defined by

$$(P_h \varphi, \chi) = (\varphi, \chi), \quad \forall\, \chi \in X_h.$$

In [13], we study the semidiscrete scheme by using FEM with $r = 1$, provided that the initial condition is regular and compatible with the boundary condition. Here

we shall briefly examine the high-order FEM, and show various convergence rates in different temporal regimes.

The time discretization is based on the so-called quadrature-based finite difference methods, which has already been intensively studied for the space-nonlocal diffusion models [8]. To this end, we partition the time interval $[0, T]$ uniformly, with grid points $t_n = n\tau$, and a time step size $\tau = T/N$. Note that the $\mathcal{G}_\delta^\tau u(t_n)$ defined in (12) is a first-order approximation of $\mathcal{G}_\delta u(t_n)$. The convergence and asymptotically compatibility has been analyzed in our preceding work [5]. Here we would like to propose time stepping schemes with higher convergence rates.

For example, we shall apply the discretization using interpolation of piecewise linear polynomials, i.e.,

$$
\begin{aligned}
&\mathcal{G}_\delta(u)(t_n)\\
=&\int_0^\delta \frac{u(t_n)-u(t_n-s)}{s}\rho_\delta(s)\,ds\\
=&\sum_{k=0}^{\delta/\tau-1}\int_{k\tau}^{(k+1)\tau}\frac{u(t_n)-u(t_n-s)}{s}\rho_\delta(s)\,ds\\
\approx&\sum_{k=0}^{\delta/\tau-1}\int_{k\tau}^{(k+1)\tau}\frac{\rho_\delta(s)}{s}\left(u(t_n)-\frac{(t_{k+1}-s)u(t_{n-k})+(s-t_k)u(t_{n-k-1})}{\tau}\right)ds\\
=:&\mathcal{G}_\delta^{\tau,2}u(t_n).
\end{aligned}
\tag{15}
$$

By Taylor's expansion, we can easily verify that the truncation error is second-order in case that the function u is regular enough with $t \in [-\delta, T]$.

Then we obtained a fully discrete scheme for approximately solving the initial-boundary value problem (8): find $U_n \in X_h$, $n = 1, 2, \ldots, N$, such that

$$
(\mathcal{G}_\delta^{\tau,2}U_n, \chi)+(\nabla U_n, \nabla\chi) = 0\ \forall\,\chi \in X_h,\ \text{with } U_{-k} = g(-t_k),\ 0 \le k \le \delta/\tau-1.
\tag{16}
$$

To test the proposed numerical scheme, we consider a one-dimensional problem in unit interval $\Omega = (0, 1)$, with the initial condition:

$$
g(x,t) := \delta(x-\frac{1}{2}), \qquad \forall\ t \in [-\delta, 0].
$$

$\delta(x-\frac{1}{2})$ denotes the Dirac-delta measure concentrated at $x = \frac{1}{2}$. Since closed-form exact solution is not available, we use very fine meshes in both space and time to compute a reference solution, i.e., $h = 10^{-3}$ and $\tau = 10^{-4}$.

First of all, we test the approximation error in space. To this end, we fix small step size $\tau = 10^{-4}$ so that the error incurred by temporal discretization is negligible. In our computation, we let $r = 4$ and $\delta = 0.2$. In Table 1, we show the space error for different α and at different T. Even though the initial data is weak, we still observe a stable convergence rate of order $O(h^{\frac{3}{2}})$ for $t \in (0, \delta)$, $O(h^{\frac{7}{2}})$ for $t \in (\delta, 2\delta)$, due to the gradual smoothing property of the nonlocal-in-time diffusion. The rigorous error analysis in different regimes is unclear, because the standard technique such as energy argument or solution operators (see e.g., [37, Chapter 3] and [20]) are not directly applicable. Besides, how to develop a high-order approximation of solution for small t is still unclear and warrants further investigation.

Next, we test the approximation error in time. In our computation, we fix $r = 4$ and $h = 1/1000$, and choose the time-independent initial condition

$$
g(x,t) := g(x) = x(1-x), \qquad \forall\ x \in (0,1)
$$

Even though the time discretization is second-order accurate (for smooth functions), we can only observe a first-order convergence in practice, cf. Table 2. This might be due to the low regularity (in the time variable) of the solution near $t = 0$, and the historical initial data is not compatible with the dynamics. To improve the convergence rate, one simple strategy is to modify the numerical scheme at the first step

$$\begin{aligned} &(\mathcal{G}_\delta^{\tau,2} U_n, \chi) + (\nabla U_n, \nabla\chi) = -\frac{1}{2}(\nabla g, \nabla\chi) \qquad \forall\, \chi \in X_h, \qquad n = 1; \\ &(\mathcal{G}_\delta^{\tau,2} U_n, \chi) + (\nabla U_n, \nabla\chi) = 0 \qquad \forall\, \chi \in X_h, \qquad n \geq 2, \end{aligned} \tag{17}$$

with $U_{-k} = P_h g(x)$ for $k = 0, \delta/\tau - 1$. This simple modification significantly improves the convergence rate for solving nonlocal-in-time diffusion (8), cf. Table 3. For $\alpha < 0$, we observe seconder-order convergence for all $t > 0$, while for $\alpha \in (0, 1)$, the numerical results show the empirical convergence rate of order $O(\tau^{2-\alpha})$ for any $t > 0$. The rigorous analysis of the modified numerical scheme (17) is challenging and differs substantially from that of its classical diffusion counterpart [24] and the time-fractional diffusion [20]. There are many interesting and open questions, e.g., how to improve the convergence rate in case of time-dependent initial condition, how to develop a second-order scheme for the case that $\alpha > 0$, whether the modified schemes are asymptotically compatible or not, etc. We plan investigate those problems in our future studies.

TABLE 1. Numerical scheme (16), error in space, with $\delta = 0.1$, $h = 1/M$, $\tau = 10^{-4}$.

α	$T\backslash M$	5	10	20	40	rate
0.25	0.1	4.99e-3	1.75e-3	6.18e-4	2.15e-4	≈ 1.52
	0.3	3.70e-6	3.03e-7	2.62e-8	2.30e-9	≈ 3.52
	0.5	8.69e-8	2.54e-9	8.25e-11	2.23e-12	≈ 5.10
−0.25	0.1	6.96e-3	2.45e-3	8.66e-4	3.02e-4	≈ 1.52
	0.3	7.93e-6	6.77e-7	5.91e-8	5.21e-9	≈ 3.52
	0.5	1.12e-7	3.32e-9	1.07e-10	2.93e-12	≈ 5.07

TABLE 2. Numerical scheme (16), error in time at $T = 0.3$, with $\delta = 0.1$, $\tau = 1/N$.

$\alpha\backslash N$	100	200	300	400	rate
0.25	2.83e-5	1.40e-5	6.98e-6	3.48e-6	≈ 1.01
0.75	3.51e-5	1.68e-5	8.11e-6	3.93e-6	≈ 1.05
−0.25	2.85e-5	1.42e-5	7.10e-6	3.54e-6	≈ 1.00
−0.75	2.92e-5	1.46e-5	7.28e-6	3.64e-6	≈ 1.00

5. Conclusion

In conclusion, through the computation of solutions and their statistics such as the mean square displacement, and through the comparisons of them with the local and fractional counterparts, this paper shows that the nonlocal-in-time diffusion with a finite memory can be a very effective model that provides an intermediate case between normal and fractional diffusions and can serve to describe effectively various processes involving crossovers of different diffusion regimes. The model is

TABLE 3. Modified scheme (17), error in time at $T = 0.3$, with $\delta = 0.1$, $\tau = 1/N$.

$\alpha \backslash N$	100	200	300	400	rate
0.25	4.18e-7	1.39e-7	4.51e-8	1.43e-8	≈ 1.65
0.75	8.62e-6	3.66e-6	1.55e-6	6.56e-7	≈ 1.24
-0.25	7.60e-8	1.83e-8	4.42e-9	1.07e-9	≈ 2.05
-0.75	1.09e-7	2.73e-8	6.82e-9	1.71e-9	≈ 2.00

showing promises in capturing experimental observations of diffusion in heterogeneous media without introducing complications and heterogeneities in the model themselves. The essence lies in the finite memory effect and how it compares with the overall dynamic history. One may naturally ask how the memory kernel and the horizon should be chosen, which becomes an interesting inverse problem to be further investigated. Other interesting issues to be studied include the connection to other models of anomalous diffusion such as those discussed in the above and in Section 1 and the replacement of normal diffusion operator (the Laplacian Δ) by nonlocal diffusion operators in the spatial directions. The latter is often associated with super-diffusion [10, 7], so that a combined nonlocal in time and space diffusion model might effectively describe sub-, normal- and super-diffusion regimes. Finally, as we discussed in last section, the distinct feature of the nonlocal model results in a lot of interesting and open questions in the simulation and related numerical analysis.

Acknowledgements

Research of Qiang Du is supported in part by NSF DMS-2012562, DMS-1937254 and NSF CNS-203894. The research of Zhi Zhou is partially supported by Hong Kong Research Grants Council (15303122) and an internal grant of Hong Kong Polytechnic University (Project ID: P0031041, Work Programme: ZZKS).

References

[1] M. Allen, L. Caffarelli, and A. Vasseur. A parabolic problem with a fractional time derivative. Archive for Rational Mechanics and Analysis, 221(2):603–630, 2016.

[2] E. Awad and R. Metzler. Crossover dynamics from superdiffusion to subdiffusion: Models and solutions. Fractional Calculus and Applied Analysis, 23(1):55–102, 2020.

[3] A. M. Berezhkovskii, L. Dagdug, and S. M. Bezrukov. Discriminating between anomalous diffusion and transient behavior in microheterogeneous environments. Biophysical Journal, 106(2):L9–L11, 2014.

[4] B. Berkowitz, J. Klafter, R. Metzler, and H. Scher. Physical pictures of transport in heterogeneous media: Advection-dispersion, random-walk, and fractional derivative formulations. Water Resources Research, 38(10):9–1, 2002.

[5] A. Chen, Q. Du, C. Li, and Z. Zhou. Asymptotically compatible schemes for space-time nonlocal diffusion equations. Chaos Solitons Fractals, 102:361–371, 2017.

[6] M. V. Chubynsky and G. W. Slater. Diffusing diffusivity: a model for anomalous, yet brownian, diffusion. Physical review letters, 113(9):098302, 2014.

[7] O. Defterli, M. D'Elia, Q. Du, M. Gunzburger, R. Lehoucq, and M. M. Meerschaert. Fractional diffusion on bounded domains. Fractional Calculus and Applied Analysis, 18(2):342–360, 2015.

[8] Q. Du. Nonlocal modeling, analysis, and computation, volume 94 of CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2019.

[9] Q. Du, M. Gunzburger, R. Lehoucq, and K. Zhou. A nonlocal vector calculus, nonlocal volume-constrained problems, and nonlocal balance laws. Mathematical Models and Methods in Applied Sciences, 23(3):493–540, 2013.

[10] Q. Du, M. Gunzburger, R. B. Lehoucq, and K. Zhou. Analysis and approximation of nonlocal diffusion problems with volume constraints. SIAM Review, 54(4):667–696, 2012.
[11] Q. Du, Y. Tao, X. Tian, and J. Yang. Robust a posteriori stress analysis for quadrature collocation approximations of nonlocal models via nonlocal gradients. Comp Meth. Applied Mech. Eng., 310:605–627, 2016.
[12] Q. Du, L. Toniazzi, and Z. Zhou. Stochastic representation of solution to nonlocal-in-time diffusion. Stochastic Process. Appl., 130(4):2058–2085, 2020.
[13] Q. Du, J. Yang, and Z. Zhou. Analysis of a nonlocal-in-time parabolic equation. Discrete Contin. Dyn. Syst. Ser. B, 22(2):339–368, 2017.
[14] W. Feller. An Introduction to Probability Theory and its Applications, volume 2. John Wiley & Sons, 2008.
[15] J. Gajda and M. Magdziarz. Fractional Fokker-Planck equation with tempered α-stable waiting times: Langevin picture and computer simulation. Phys. Rev. E (3), 82(1):011117, 6, 2010.
[16] D. S. Grebenkov. A unifying approach to first-passage time distributions in diffusing diffusivity and switching diffusion models. Journal of Physics A: Mathematical and Theoretical, 52(17):174001, 2019.
[17] W. He, H. Song, Y. Su, L. Geng, B. Ackerson, H. Peng, and P. Tong. Dynamic heterogeneity and non-Gaussian statistics for acetylcholine receptors on live cell membrane. Nature communications, 7, 2016.
[18] R. Jain and K. Sebastian. Diffusing diffusivity: a new derivation and comparison with simulations. Journal of Chemical Sciences, 129(7):929–937, 2017.
[19] J.-H. Jeon, H. M.-S. Monne, M. Javanainen, and R. Metzler. Anomalous diffusion of phospholipids and cholesterols in a lipid bilayer and its origins. Physical Review Letters, 109(18):188103, 2012.
[20] B. Jin, R. Lazarov, and Z. Zhou. Numerical methods for time-fractional evolution equations with nonsmooth data: A concise overview. Comput. Methods Appl. Mech. Engrg., 346:332–358, 2019.
[21] B. Jin and Z. Zhou. Numerical Treatment and Analysis of Time-Fractional Evolution Equations. Springer Cham, 2023.
[22] A. Kilbas, H. Srivastava, and J. Trujillo. Theory and Applications of Fractional Differential Equations. Elsevier, Amsterdam, 2006.
[23] N. Korabel and E. Barkai. Paradoxes of subdiffusive infiltration in disordered systems. Physical Review Letters, 104(17):170603, 2010.
[24] B. Li, K. Wang, and Z. Zhou. Long-time accurate symmetrized implicit-explicit BDF methods for a class of parabolic equations with non-self-adjoint operators. SIAM J. Numer. Anal., 58(1):189–210, 2020.
[25] F. Mainardi, Y. Luchko, and G. Pagnini. The fundamental solution of the space-time fractional diffusion equation. Fractional Calculus and Applied Analysis, 4(2):153–192, 2001.
[26] S. A. McKinley, L. Yao, and M. G. Forest. Transient anomalous diffusion of tracer particles in soft matter. Journal of Rheology, 53(6):1487–1506, 2009.
[27] T. Mengesha and Q. Du. Characterization of function spaces of vector fields and an application in nonlinear peridynamics. Nonlinear Analysis A: Theory, Methods and Applications, 140:82–111, 2016.
[28] R. Metzler. Brownian motion and beyond: first-passage, power spectrum, non-gaussianity, and anomalous diffusion. arXiv preprint arXiv:1908.06233, 2019.
[29] R. Metzler and J. Klafter. The random walk's guide to anomalous diffusion: a fractional dynamics approach. Physics Reports, 339(1):1–77, 2000.
[30] R. Metzler and J. Klafter. The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics. Journal of Physics A: Mathematical and General, 37(31):R161, 2004.
[31] F. Sabzikar, M. M. Meerschaert, and J. Chen. Tempered fractional calculus. J. Comput. Phys., 293:14–28, 2015.
[32] M. J. Saxton. Wanted: a positive control for anomalous subdiffusion. Biophysical Journal, 103(12):2411–2422, 2012.
[33] I. M. Sokolov. Models of anomalous diffusion in crowded environments. Soft Matter, 8(35):9043–9052, 2012.
[34] V. Sposini, A. V. Chechkin, F. Seno, G. Pagnini, and R. Metzler. Random diffusivity from stochastic equations: comparison of two models for brownian yet non-gaussian diffusion. New Journal of Physics, 20(4):043044, 2018.

[35] H. Sun, Z. Li, Y. Zhang, and W. Chen. Fractional and fractal derivative models for transient anomalous diffusion: Model comparison. Chaos, Solitons and Fractals, 102:346–353, 2017.
[36] P. Tan, Y. Liang, Q. Xu, E. Mamontov, J. Li, X. Xing, and L. Hong. Gradual crossover from subdiffusion to normal diffusion: a many-body effect in protein surface water. Physical review letters, 120(24):248101, 2018.
[37] V. Thomée. Galerkin Finite Element Methods for Parabolic Problems, volume 25 of Springer Series in Computational Mathematics. Springer-Verlag, Berlin, 2006.
[38] H. Wang and X. Zheng. Wellposedness and regularity of the variable-order time-fractional diffusion equations. J. Math. Anal. Appl., 475(2):1778–1802, 2019.
[39] X. Wu, W. Deng, and E. Barkai. Tempered fractional Feynman-Kac equation: theory and examples. Phys. Rev. E, 93(3):032151, 15, 2016.
[40] X. Zheng, J. Cheng, and H. Wang. Uniqueness of determining the variable fractional order in variable-order time-fractional diffusion equations. Inverse Problems, 35(12):125002, 11, 2019.

Department of Applied Physics and Applied Mathematics, Columbia University, New York 10027, USA
E-mail: qd2125@columbia.edu

Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong SAR, China
E-mail: zhizhou@polyu.edu.hk