ON AN OPTIMAL CONTROL PROBLEM OF
TIME-FRACTIONAL ADVECTION-DIFFUSION EQUATION

QING TANG∗
China university of Geosciences, Wuhan
Hubei province, China

Abstract. We consider an optimal control problem of an advection-diffusion equation with Caputo time-fractional derivative. By convex duality method we obtain as optimality condition a forward-backward coupled system. We then prove the existence of a solution to this coupled system using Schauder fixed point theorem. The uniqueness of the solution is also established under certain monotonicity condition on the cost functional.

1. Introduction. In this paper we study the following time-fractional advection-diffusion (AD) equation

\[
\begin{cases}
\partial_{(0,t]}^\beta m - \Delta m + \text{div}(mv(t,x)) = 0, & (t,x) \in (0,T) \times \mathbb{T}^d \\
m(0,x) = m_0(x).
\end{cases}
\]

(1)

Throughout this paper, we always assume that \( \beta \in (0,1] \). We consider optimal control problem with the cost functional:

\[
\inf_{v,m} J_{\text{AD}} = \inf_{v,m} \int_0^T \int_{\mathbb{T}^d} \frac{1}{2} |v|^2 m + F(x,m) \, dxdt + \int_{\mathbb{T}^d} u_T(x)(t^{1-\beta})^\beta(T)dx.
\]

As optimality condition we obtain the backward-forward system of coupled equations, \( F \) is the primitive of \( f \) with respect to \( m \):

\[
\begin{cases}
\partial_{[t,T]}^\beta u - \Delta u + \frac{1}{2} |\nabla u|^2 = f(x,m), & (t,x) \in (0,T) \times \mathbb{T}^d \\
\partial_{(0,t]}^\beta m - \Delta m + \text{div}(-m\nabla u) = 0,
\end{cases}
\]

(2)

Here \( \partial_{[t,T]}^\beta \) and \( \partial_{(0,t]}^\beta \) denote the backward and forward Caputo fractional derivatives. To avoid both compactness and boundary conditions issues, instead working on a domain or \( \mathbb{R}^d \), we will consider the case of the flat torus \( \mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d \). This choice is standard for works in forward-backward systems, e.g. [14]. The term \( f \) associates to a probability density \( m \) a real valued function \( f(x,m) \). \( v = v(t,x) \) is time dependent. \( \nabla \) and \( \Delta \) denote the gradient and Laplacian operators.

In [30], Li and Liu studied the existence, uniqueness and regularity of solution of fractional advection-diffusion equation (1) as a preliminary step to the study of a time-fractional Keller-Segel system in domain \( \mathbb{R}^2 \). Recently there has been growing interest in studying the non-local in time equations with similar structure. When

2010 Mathematics Subject Classification. Primary: 35R11, 60H05, 26A33, 40L20.
Key words and phrases. Optimal control, Caputo derivative, fractional Fokker-Planck equation, fractional Hamilton-Jacobi equation, variational Mean Field Games.

∗ Corresponding author: Qing Tang.
the drift term \( v \) does not depend on time, for example when \( v(x) = -\nabla V(x) \) with \( V(x) \) being a potential, the equation

\[
\begin{align*}
\partial_t m - \Delta m + \text{div}( -m \nabla V(x)) = 0, \quad (t, x) \in (0, T) \times \mathbb{T}^d \\
m(0, x) = m_0(x)
\end{align*}
\]

is called a time-fractional Fokker Planck equation (FP equation), or Kolmogorov forward equation. Equation (3) may be used to describe the flow in time of probability density in subdiffusion regimes, i.e. particles are subjected to long-tailed random waiting time. Kemppainen and Zacher [24] studied the long-time behavior of a time-fractional FP equation via the entropy method. Hahn, Kobayashi and Umarov [20] derived this equation from a model of stochastic differential equations driven by time-changed Lévy processes.

Very important applications in physics and finance, we refer to [8, 33] for general theory of anomalous diffusion. This type of time-fractional parabolic equations has been intensively investigated in recent works, e.g. [30, 2, 37, 38, 27]. Optimal control problems of time-fractional diffusion type equations and numerical schemes have been considered by Antil et al. in [4] and by Annunziato et al. in [3].

Subdiffusion processes deviate from classical Gaussian process in that the motion of particles are interrupted by long sojourns, possibly due to trapping effects. It has very important applications in physics and finance, we refer to [8, 33] for general theory of anomalous diffusion.

The adjoint equation we obtained in the optimality system (2) is a backward Hamilton-Jacobi equation with Caputo derivative:

\[
\begin{align*}
\partial_{[t,T]} u - \Delta u + \frac{1}{2} |\nabla u|^2 &= f(x, m), \quad (t, x) \in (0, T) \times \mathbb{T}^d \\
u(T, x) &= u_T(x).
\end{align*}
\]

We note that by change of variable (from \( t \) to \( T - t \)), it may be transformed into

\[
\begin{align*}
\partial_{[0,T]} u - \Delta u + \frac{1}{2} |\nabla u|^2 &= f(x, m), \quad (t, x) \in (0, T) \times \mathbb{T}^d \\
u(0, x) &= u_T(x).
\end{align*}
\]

In the case that \( f \) is sufficiently regular and does not depend on density \( m \), i.e. \( f(x, m) = f(t, x) \in W^{1,\infty}(0, T; W^{1,\infty}(\mathbb{T}^d)) \), the classical solution to the forward fractional HJ equation has been obtained in [25, 26, 22] by Kolokoltsov et al. The
viscosity solution theory for time-fractional PDEs was also developed in some recent works, e.g. [18, 36]. In [10], Camilli, Maio and Iacomini introduced a Hopf-Lax formula for the solution of a fractional Hamilton Jacobi equation.

When $\beta = 1$ the system (2) becomes the classical Lasry-Lions mean field game system:

$$\begin{align}
-\partial_t u - \Delta u + \frac{1}{2} |\nabla u|^2 &= f(x, m), \quad (t, x) \in (0, T) \times T^d \\
\partial_t m - \Delta m + \text{div}(-m \nabla u) &= 0, \\
m(0, x) &= m_0(x), \quad u(T, x) = u_T(x).
\end{align}$$

The theory of Mean Field Games (MFG, in short) is a branch of Dynamic Games which aims at modeling and analyzing complex decision processes involving a large number of indistinguishable rational agents who have individually a very small influence on the overall system and are, on the other hand, influenced by the distribution of the other agents (Achdou et al. [1]). The theory of Mean Field Games originated from the works of Lasry and Lions [28, 29] and independently started by Huang, Caines and Malhamé [23]. For a general introduction of MFG we refer to [12, 17]. The system (10) may be derived via a convex optimization problem of the Fokker-Planck equation, this is known as the theory of variational mean field games [13, 14]. This approach is based on the dynamic formulation of the Monge-Kantorovich optimal transport problem proposed by Benamou and Brenier [5]. Recent advances in this direction include [6, 32, 7]. An application of variational MFG to nonlinear mobilities in pedestrian dynamics can be found in [9].

Camilli and Maio [11] considered a time-fractional MFG system:

$$\begin{align}
-\partial_t u + D^{1-\beta}_{[t,T]}[-\Delta u + \frac{1}{2} |\nabla u|^2] &= f(x, m), \quad (t, x) \in (0, T) \times \mathbb{R}^d \\
\partial_t m - [\Delta + \text{div}(\nabla u)](D^{1-\beta}_{[0,t]} m) &= 0, \\
m(0, x) &= m_0(x), \quad u(T, x) = u_T(x),
\end{align}$$

where $D^{1-\beta}_{[t,T]}$ and $D^{1-\beta}_{[0,t]}$ denote the backward and forward Riemann-Liouville fractional derivatives. Later, Tang and Camilli in [35] have derived the system (11) from a variational MFG approach. It is important to note the relation and differences between the two systems (2) and (11). The system (2) is the variant of the classical MFG system with Caputo time derivatives. It has the advantage of being more trackable via PDE methods, and may be naturally extended to include space non-local diffusion (by using the fractional Laplacian instead of Laplacian). The positivity and mass preservation of solutions to equation (1) has been proven in [30]. This equation itself is also instrumental in modeling of anomalous diffusion and aggregation processes, e.g. chemotaxis of bacteria. However, it has the drawback that the game theory interpretation is not clear, as the fractional HJ equation has not been derived via dynamical programming from a single agent point of view, but only as the adjoint equation in a PDE control problem. In comparison, in system (11), Camilli and Maio in [11] have proved it to be a bona fide MFG system. Therefore we believe the two fractional parabolic systems are of independent interests and both worthy of study.

The following assumptions and notations are supposed to hold throughout the rest of the paper.

(H1) The coupling $f : T^d \times [0, +\infty) \to \mathbb{R}$ is continuous in both variables. Moreover the following normalization condition holds:

$$f(x, 0) = 0, \forall x \in T^d.$$
(H2) \( f(x, m) \) is increasing with respect to \( m \), i.e., \( \forall m_1, m_2 \in C([0, T]; \mathcal{P}_1), \forall t \in [0, T], \mathcal{P}_1 \) denotes the space of probability measures on \( \mathbb{T}^d \) endowed with the Monge-Kantorovich distance,

\[
\int_{\mathbb{T}^d} (f(x, m_1) - f(x, m_2))(m_1 - m_2) \, dx \geq 0.
\]

(H3) \( u_T(x) : \mathbb{T}^d \to \mathbb{R} \) is of class \( C^2 \), while \( m_0(x) : \mathbb{T}^d \to \mathbb{R} \) is of class \( C^1 \) positive density, i.e., \( m_0(x) > 0 \) and \( \int_{\mathbb{T}^d} m_0(x) \, dx = 1 \).

Let us set:

\[
F(x, m) = \begin{cases} 
\int_0^m f(x, \tau) \, d\tau, & \text{if } m \geq 0, \\
+\infty & \text{otherwise}.
\end{cases}
\]

From conditions (H1) and (H2) it follows that \( F(x, m) \) is convex with regard to \( m \).

** PLAN OF THE PAPER.** The paper is organized as follows. In Section 2, we review some basic facts about the fractional calculus, introduce the notion of weak solution to the time-fractional advection-diffusion equation and prove the existence and uniqueness of the solution. In Section 3, we derive the coupled fractional forward-backward system (2) as the optimality condition of control problems driven by the fractional advection-diffusion equation and the equivalence with a control problem of the fractional HJ equation via Fenchel-Rockefeller duality theorem. We also give verification results regarding the equivalence between solutions of coupled fractional system and the two optimization problems. Finally, we discuss some interesting directions of investigation for future work.

2. **Fractional calculus and advection-diffusion equation.**

2.1. **Fractional calculus.** The idea of defining a derivative of fractional order \( \frac{1}{2} \) for example) dates back to Leibniz. This problem has also been considered by Riemann and Liouville among others in 19th century. The nonlocal operators under the name Riemann-Liouville derivative and integral are the most important definitions in this subject to this day. In the first half of 20th century advances has been made by Hardy and Littlewood \cite{21}. We refer to Samko, Kilbas and Marichev \cite{34} for a comprehensive account of the theory. We start with a brief introduction to some definitions and basic results in fractional calculus. We start from a formal level and assume functions \( \phi(t), \kappa(t) \in C^1(\mathbb{R}) \).

The forward and backward Riemann-Liouville fractional integrals are defined by

\[
\begin{align*}
I_{(0,t)}^{\beta} \phi(t) &:= \frac{1}{\Gamma(\beta)} \int_0^t \phi(\tau) \frac{1}{(t-\tau)^{1-\beta}} \, d\tau, \\
I_{(t,T)}^{\beta} \phi(t) &:= \frac{1}{\Gamma(\beta)} \int_t^T \phi(\tau) \frac{1}{(\tau-t)^{1-\beta}} \, d\tau.
\end{align*}
\]

The forward Riemann-Liouville and Caputo derivatives are defined by

\[
\begin{align*}
D_{(0,t)}^{\beta} \phi(t) &:= \frac{d}{dt} \left[ I_{(0,t)}^{1-\beta} \phi(t) \right] = \frac{1}{\Gamma(1-\beta)} \int_0^t \frac{d\phi}{dt}(\tau) \frac{1}{(t-\tau)^{\beta}} \, d\tau, \\
\partial_{(0,t)}^{\beta} \phi(t) &:= \frac{d}{dt} \left[ I_{(0,t)}^{1-\beta} \phi(t) \right] = \frac{1}{\Gamma(1-\beta)} \int_0^t \frac{d\phi}{dt}(\tau) \frac{1}{(t-\tau)^{\beta}} \, d\tau,
\end{align*}
\]

while the backward Riemann-Liouville and Caputo derivatives are defined by

\[
\begin{align*}
D_{(t,T)}^{\beta} \phi(t) &:= -\frac{d}{dt} \left[ I_{(t,T)}^{1-\beta} \phi(t) \right] = -\frac{1}{\Gamma(1-\beta)} \int_t^T \frac{d\phi}{dt}(\tau) \frac{1}{(\tau-t)^{\beta}} \, d\tau,
\end{align*}
\]
\[ \partial_{[t,T]}^\beta \phi(t) := -I_{[t,T]}^{1-\beta} \left[ \frac{d\phi(t)}{dt} \right] = -\frac{1}{\Gamma(1-\beta)} \int_t^T \frac{d\phi(\tau)}{d\tau} \frac{1}{(t-\tau)^\beta} d\tau. \]

For \( \beta \to 1 \) the forward Riemann-Liouville and Caputo derivatives of \( \phi \) converge to the classical derivative \( \frac{d\phi}{dt} \), meanwhile, the backward derivatives converge to \( -\frac{d\phi}{dt} \).

It should be noticed that the Riemann-Liouville derivative of a constant is not 0:
\[
D_{(0,a)}^{1-\beta} 1 = \frac{1}{\Gamma(\beta)} \frac{d}{da} \int_0^a \frac{d\phi(\tau)}{d\tau} (1-\beta)^{-1} d\tau = \frac{t^{\beta-1}}{\Gamma(\beta)}.
\]

We note the following relations between fractional integral and Caputo derivative, which may be regarded as the fractional version of fundamental theorem of calculus:
\[
\phi(t) = \phi(0) + I_{(0,a)}^\beta (\partial_{(0,a)}^\beta \phi) = \phi(0) + \frac{1}{\Gamma(\beta)} \int_0^a (t-\tau)^{\beta-1} \partial_{(0,a)}^\beta \phi d\tau.
\] (12)

The following fractional version to the classic Gronwall lemma will be used in priori estimates for solution to fractional PDEs.

**Lemma 2.1.** Let \( \phi(t) \in C^1([0,T]) \), and constant \( \lambda \neq 0 \),
\[ \partial_{(0,t]}^\beta \phi \leq \lambda \phi(t), \]
then
\[ \phi(t) \leq \phi(0) \sum_{k=0}^{\infty} \lambda^k \frac{t^{\beta k}}{\Gamma(1+\beta k)}. \]

**Proof.** This lemma is just a particular case of Proposition 6.6 of [27], using (12) and the fact that
\[
\sum_{k=0}^{\infty} \lambda^k I_{(0,a)}^{\beta k} (\phi(0)) = \phi(0) \sum_{k=0}^{\infty} \lambda^k (I_{(0,a)}^{\beta k} \cdot 1),
\]
and
\[
I_{(0,a)}^{\beta k} \cdot 1 = \frac{1}{\Gamma(\beta k)} \frac{t^{\beta k}}{\beta k} = \frac{t^{\beta k}}{\Gamma(1+\beta k)}.
\]

Moreover, we have the following result regarding the fractional integration by parts with Caputo derivative ([4], Lemma 3):

**Lemma 2.2.** Let \( \phi(t), \kappa(t) \in C^1([0,T]) \), then
\[
\int_0^T \partial_{(0,t]}^{\beta} \phi(t) \kappa(t) dt + \phi(0)(I_{(0,t]}^{1-\beta} \kappa)(0) = \int_0^T \phi(t)(\partial_{[t,T]}^{\beta} \kappa(t)) dt + \kappa(T)(I_{(0,a)}^{1-\beta} \phi)(T).
\] (13)

This is equivalent to
\[
\int_0^T (\partial_{(0,t]}^{\beta} \phi)(\kappa(t) - \kappa(T)) dt = \int_0^T (\partial_{[t,T]}^{\beta} \kappa)(\phi(t) - \phi(0)) dt.
\] (14)

In order to study partial differential equations with fractional time derivatives, one needs to construct the fractional calculus for functions valued in general Banach spaces. We refer readers to [30] by Li and Liu for this theory. Denote by \( B \) a Banach space, we introduce the following sets:
\[
\mathcal{G} := \{ v | v : C^\infty_c((-\infty, T); \mathbb{R}) \to B \text{ is a bounded linear operator} \}. 
\]
We are motivated by the fractional integration by parts Lemma (2.2) to define the weak Caputo derivative in the sense of distributions:

**Definition 2.3.** Let \( u \in L^1([0,T);B) \) and \( u_0 \in B \). We define the weak Caputo derivative of \( u(t,x) \) at time \( t \), associated with initial data \( u_0 \), to be \( \partial^\beta_{(0,t)} u \in D' \) such that for any test function \( \varphi \in C_c^\infty([0,T);B) \),

\[
\int_0^T \partial^\beta_{(0,t)} u \varphi dt = \int_0^T u \partial^\beta_{(t,T)} \varphi dt - u_0(\partial^\beta_{(0,T)} \varphi)(0).
\]

The following result ([30], Lemma 3.1) is due to Hardy and Littlewood (Theorem 12 of [21]). An alternative proof has recently been obtained by Kubica and Yamamoto ([27], Proposition 6.7).

**Lemma 2.4.** Let \( B \) be a Banach space and \( T > 0 \), \( \beta \in (0,1) \). Suppose the Caputo derivative \( \partial^\beta_{(0,t)} u \in L^p([0,T);B) \) and \( p > \frac{1}{\beta} \), then \( u \) is continuous on \([0,T]\) such that

\[
\|u(t+h) - u(t)\|_B \leq C_1 h^{\beta-1/p},
\]

for \( 0 \leq t < t+h \leq T \) and \( C_1 \) is independent of \( t \).

### 2.2. Time-fractional advection diffusion equation

We define the weak formulation to the fractional advection-diffusion equation:

**Definition 2.5.** Let \( \beta \in (0,1) \). We say that \( m \) is a weak solution to the fractional advection-diffusion equation (1), if for any test function \( \varphi(t,x) \in C_c^\infty([0,T) \times \mathbb{T}^d;\mathbb{R}^d) \):

\[
\int_0^T \int_{\mathbb{T}^d} \partial^\beta_{(t,T)} \varphi(m(\tau,x) - m_0) dx d\tau - T \int_{\mathbb{T}^d} m \Delta \varphi dxdt - T \int_{\mathbb{T}^d} \nabla \varphi dxdt = 0.
\]

We also note the notion of mild solution to (1). Introduce the Mittag-Leffler function:

\[
E_\beta(z) = \sum_{k=0}^\infty \frac{z^k}{\Gamma(k\beta + 1)},
\]

by taking the Laplace transform one can obtain the fractional (non-Markovian) version of Duhamel’s principle for solution to the equation (1):

\[
m(t,x) = E_\beta(-t^\beta(-\Delta))m_0 + \beta \int_0^T \tau^{\beta-1} E'_\beta(-\tau^\beta(-\Delta))(-\text{div}(mv)|_{t-\tau}) d\tau,
\]

where

\[
\text{div}(mv)|_{t-\tau} = \text{div}(m(t-\tau,x)v(t-\tau,x)).
\]

**Definition 2.6.** Suppose \( \mathbb{B} \) is a Banach space in space and time. If \( m \in \mathbb{B} \) satisfies (15), then we say that \( m \) is a mild solution in \( \mathbb{B} \).

**Lemma 2.7.** (Lemma 5.3 of [30]) Suppose \( v(t,x) \) is smooth and all derivatives are bounded. Then:

(i.) If \( m_0 \in L^1(\mathbb{T}^d) \cap H^2(\mathbb{T}^d) \), then \( \forall T > 0, \) (1) has a unique mild solution in \( C([0,T];H^2(\mathbb{T}^d)) \).

(ii.) For the unique mild solution in (i), \( \forall T > 0, \)

\[
m \in C^\beta([0,T];H^2(\mathbb{T}^d)) \cap C^\infty((0,T);H^2(\mathbb{T}^d)).
\]

The mild solution is a strong solution in \( C([0,T];L^2(\mathbb{T}^d)) \).
Proof. We start with some basic a priori estimates for equation (1). We will find a constant $c_m$ (iii.) If $m_0 \in L^1(\mathbb{T}^d) \cap H^1(\mathbb{T}^d)$ and $m_0 \geq 0$, then $m(t, x) \geq 0$, and $\forall t \in [0, T]$ 
\[ \int_{\mathbb{T}^d} m(t, x) dx = \int_{\mathbb{T}^d} m_0 dx. \]

Remark 1. The Lemma 2.7 was proven in [30] for space domain $\mathbb{R}^2$, but the proof can be repeated for $\mathbb{T}^d$ without essential changes.

Throughout the paper we denote by $\| \cdot \|_p$ the norm of Lebesgue space $L^p(\mathbb{T}^d)$ with $p \in [1, +\infty)$.

Theorem 2.8. Suppose $m_0 \in L^1(\mathbb{T}^d) \cap H^2(\mathbb{T}^d)$ and $m_0 \geq 0$, $v \in L^{\infty}([0, T]; L^\infty(\mathbb{T}^d))$, then there is a unique weak solution $m$ to equation (1) in the space $\Xi \cap L^\infty([0, T]; L^2(\mathbb{T}^d))$.

Moreover, $\forall t \in [0, T]$, we have positivity and mass conservation of the solution, i.e. $m(t, x) \geq 0$ and
\[ \int_{\mathbb{T}^d} m(t, x) dx = \int_{\mathbb{T}^d} m_0 dx. \]

If $\beta > \frac{1}{2}$, we have solution $m \in C([0, T]; H^{-1}(\mathbb{T}^d))$.

Proof. We start with some basic a priori estimates for equation (1). We will find a constant $c_0 > 0$ such that uniformly $\forall t \in [0, T]$:
\[ \|v(t, x)\|_{\infty} \leq c_0. \]

From Poincare inequality there exists a constant $c_2 > 0$ such that $\|m(t, \cdot)\|_2 \leq c_2\|\nabla m(t, \cdot)\|_2$. Then $\forall t \in (0, T)$,
\[ \partial^\beta_{(0, t)} \left( \frac{1}{2} \|m\|_2^2 \right) \leq \langle m, \partial^\beta_{(0, t)} m \rangle \]
\[ = \langle m, \Delta m - \text{div}(vm) \rangle \]
\[ = -\int_{\mathbb{T}^d} |\nabla m|^2 dx + \int_{\mathbb{T}^d} \nabla m \cdot (vm) dx \]
\[ \leq -\int_{\mathbb{T}^d} |\nabla m|^2 dx + c_0 \int_{\mathbb{T}^d} |m\nabla m| dx \]
\[ \leq c_0 \left( \frac{1}{c_0} \int_{\mathbb{T}^d} |\nabla m|^2 dx + \frac{1}{c_0} \int_{\mathbb{T}^d} |\nabla m|^2 dx + \frac{c_0}{4} \int_{\mathbb{T}^d} m^2 dx \right) \]
\[ = \frac{c_0^2}{4} \|m\|_2^2. \]

The first inequality is due to the Proposition 2.2 of [30]. We then apply Lemma 2.1 and obtain
\[ \|m(t, x)\|_2^2 \leq \left( \sum_{k=0}^{\infty} \frac{c_0^2}{2} \frac{t^{\beta k}}{\Gamma(1 + \beta k)} \right) \|m_0(x)\|_2^2 \leq \left( \sum_{k=0}^{\infty} \frac{c_0^2 T^\beta}{2} \frac{1}{\Gamma(1 + \beta k)} \right) \|m_0(x)\|_2^2. \]

This series is indeed convergent by d’Alembert criterion. Recall that for $\beta \in (0, 1]$, $1 + \beta k > 1$, $\lim_{x \to \infty} \frac{\Gamma(x + \alpha)}{\Gamma(x) x^\alpha} = 1$.
so that we have
\[ \lim_{k \to \infty} \frac{c_0^2}{2} \frac{\Gamma(1 + \beta k)}{\Gamma(1 + \beta (k + 1))} \leq \frac{c_0^2}{2} \frac{\Gamma(1 + \beta k)}{\Gamma(1 + \beta k + \beta)} = \lim_{k \to \infty} \frac{c_0^2}{2} \frac{1}{(1 + \beta k)^\beta} = 0. \]

Thus we may conclude that \( m \in L^\infty([0, T]; L^2(\mathbb{T}^d)) \). Moreover from the calculations above we have
\[ \partial^{\beta}_{(0, t)} (\frac{1}{2} \| m \|_2^2) + \int_{\mathbb{T}^d} |\nabla m|^2 \, dx \leq c_0 \int_{\mathbb{T}^d} |m \nabla m| \, dx \leq c_0 (\frac{1}{2c_0} \int_{\mathbb{T}^d} |\nabla m|^2 \, dx + \frac{c_0}{2} \int_{\mathbb{T}^d} m^2 \, dx), \]
so that
\[ \partial^{\beta}_{(0, t)} (\| m \|_2^2) + \int_{\mathbb{T}^d} |\nabla m|^2 \, dx \leq c_0^2 \int_{\mathbb{T}^d} m^2 \, dx. \] (17)
Integrate both sides of this inequality on \([0, T]\) we obtain:
\[ I^{\beta}_{(0, t)} (\| m \|_2^2) + \int_0^T \int_{\mathbb{T}^d} |\nabla m|^2 \, dx \, dt \leq c_0^2 \int_0^T \int_{\mathbb{T}^d} m^2 \, dx \, dt \]
\[ \leq c_0^2 T \left( \sum_{k=0}^\infty \frac{c_0^2 \Gamma(1 + \beta k)}{2} \frac{1}{\Gamma(1 + \beta k)^\beta} \right) \| m_0(x) \|_2. \]

Thus
\[ \| m \|_{L^2([0, T]; H^1(\mathbb{T}^d))} = \| m \|_{L^2([0, T]; L^2(\mathbb{T}^d))} + \| \nabla m \|_{L^2([0, T]; L^2(\mathbb{T}^d))} \]
\[ \leq (1 + c_2) \| \nabla m \|_{L^2([0, T]; L^2(\mathbb{T}^d))} \]
\[ \leq (1 + c_2) (c_0^2 T)^{\frac{1}{2}} \left( \sum_{k=0}^\infty \frac{c_0^2 \Gamma(1 + \beta k)}{2} \frac{1}{\Gamma(1 + \beta k)^\beta} \right)^{\frac{1}{2}} \| m_0(x) \|_2. \]
We now turn to the estimates on \( \partial^{\beta}_{(0, t)} m \). For all test function
\[ \varphi(t, x) \in L^2([0, T]; H^1(\mathbb{T}^d)), \]
\[ \langle \partial^{\beta}_{(0, t)} m, \varphi \rangle = - \int_0^T \int_{\mathbb{T}^d} \nabla m \cdot \nabla \phi \, dx \, dt + \int_0^T \int_{\mathbb{T}^d} \nabla \varphi \cdot (v m) \, dx \, dt \]
\[ \leq \frac{1}{2} \int_{\mathbb{T}^d} (\| \nabla m \|_2^2 + c_0 \| m \|_2^2) \, dt + c_0 \int_0^T \| \nabla \varphi \|_2^2 \, dt \]
\[ < +\infty, \]
therefore we have
\[ \| \partial^{\beta}_{(0, t)} m \|_{L^2([0, T]; H^{-1}(\mathbb{T}^d))} = \sup_{\| \varphi \|_1 = 1} \langle \partial^{\beta}_{(0, t)} m, \varphi \rangle < +\infty, \]
where \( \| \varphi \| := \| \varphi \|_{L^2([0, T]; H^{-1}(\mathbb{T}^d))}. \) Let \( \chi(x) \in C_0^\infty(\mathbb{R}^{d+1}) \) is chose to be compacted supported, \( \chi(t, x) \geq 0 \) and
\[ \int_{\mathbb{R}^{d+1}} \chi(t, x) \, dx \, dt = 1. \]
Introduce the mollification
\[ v_\epsilon = \frac{1}{\epsilon^{d+1}} \int_{\mathbb{R}^{d+1}} \chi(\frac{t-s}{\epsilon}, \frac{x-y}{\epsilon}) v(s, y) \, ds \, dy. \]
We now consider the mollified equation
\[
\begin{aligned}
\partial_{(0,t)}^\beta m - \Delta m + \text{div}(mv_\epsilon) &= 0, \quad (t,x) \in (0,T) \times \mathbb{T}^d, \\
m(0,x) &= m_0(x).
\end{aligned}
\]
(18)

Using Lemma 2.7 we see that there exists a unique mild solution \(m_\epsilon\) such that
\[
m_\epsilon \in C^0([0,T]; H^2(\mathbb{T}^d)) \cap C^\infty((0,T); H^2(\mathbb{T}^d)).
\]

We note that the previous a priori estimates results we obtained for \(m\) holds for \(m_\epsilon\) uniformly in \(\epsilon\) such that
\[
m_\epsilon \in L^2([0,T]; H^1(\mathbb{T}^d)), \quad \partial_{(0,t)}^\beta m_\epsilon \in L^2([0,T]; H^{-1}(\mathbb{T}^d)), \quad \sup_{t \in (0,T)} \int_{(0,t)}^{1-\beta}(\|m_\epsilon\|_2^2) \leq C_1.
\]

Passing to the limit \(\epsilon \to 0\), since \(L^2([0,T]; H^1(\mathbb{T}^d))\) is a reflexive Banach spaces (so is its dual space \(L^2([0,T]; H^{-1}(\mathbb{T}^d))\), of course), we may construct subsequence \(m_{\epsilon k}\) such that \(m_{\epsilon k}\) converges weakly in \(L^2([0,T]; H^2(\mathbb{T}^d))\) in the weak* topology, \(m_{\epsilon k}\) converges weakly in \(L^2([0,T]; H^2(\mathbb{T}^d))\) and \(\partial_{(0,t)}^\beta m_{\epsilon k}\) converges weakly in \(L^2([0,T]; H^{-1}(\mathbb{T}^d))\). More precisely, \(\forall \varphi \in \Xi^*\) (and therefore \(\forall \varphi(t,x) \in C^\infty_c([0,T] \times \mathbb{T}^d; \mathbb{R}^d)\) via a density argument), for \(\epsilon_k \to 0\):
\[
\begin{aligned}
\int_0^T \int_{\mathbb{T}^d} \varphi \partial_{(0,t)}^\beta m_{\epsilon_k} dxdt &\to \int_0^T \int_{\mathbb{T}^d} \varphi \partial_{(0,t)}^\beta m_\epsilon dxdt, \\
\int_0^T \int_{\mathbb{T}^d} \varphi \Delta_{\epsilon_k} m dxdt &= -\int_0^T \int_{\mathbb{T}^d} \nabla \varphi \cdot \nabla m_{\epsilon_k} dxdt \\
&\to -\int_0^T \int_{\mathbb{T}^d} \nabla \varphi \cdot \nabla m_{\epsilon_k} dxdt = \int_0^T \int_{\mathbb{T}^d} \varphi \Delta_{\epsilon_k} m dxdt.
\end{aligned}
\]

Since \(v_\epsilon\) convergent strongly to \(v\), using classical diagonal argument we can subtract a further subsequence \((m_{\epsilon_k}, v_{\epsilon_k})\) such that, for \(\epsilon_k \to 0\):
\[
\begin{aligned}
\int_0^T \int_{\mathbb{T}^d} \varphi \text{div}(v_{\epsilon_k} m_{\epsilon_k}) dxdt &= \int_0^T \int_{\mathbb{T}^d} \nabla \varphi \cdot (v_{\epsilon_k} m_{\epsilon_k}) dxdt \\
&\to \int_0^T \int_{\mathbb{T}^d} \nabla \varphi \cdot (v m) dxdt = \int_0^T \int_{\mathbb{T}^d} \varphi \text{div}(vm) dxdt.
\end{aligned}
\]

Uniqueness follows analogously via a priori estimates. Suppose there exist two solutions \(m_1\) and \(m_2\) such that \(m_1(0,x) = m_2(0,x)\). Due to the linear structure of the equation it is easy to obtain from the above calculations
\[
\begin{aligned}
&\int_{(0,t)}^{1-\beta}(\|m_1 - m_2\|_2^2) + \int_0^T \int_{\mathbb{T}^d} |\nabla m_1 - \nabla m_2|^2 dxdt \\
\leq& c_2 T \sum_{k=0}^\infty \left(\frac{2^\beta}{2}\right)^k \frac{1}{\Gamma(1+\beta k)} \|m_1(0,x) - m_2(0,x)\|_2^2 = 0.
\end{aligned}
\]

Hence the solution is unique in \(\Xi\).

From Lemma 2.7 we have \(\forall t \in [0,T], m_\epsilon(t,x) \geq 0\) and \(\int_{\mathbb{T}^d} m_\epsilon(t,x) dx = 1\).

Passing to the limit \(\epsilon \to 0\), since \(m_\epsilon(t,x)\) converges locally uniformly to \(m(t,x)\), we have \(m(t,x) \geq 0\) and \(\int_{\mathbb{T}^d} m(t,x) dx = 1\).

Finally, if \(\beta > \frac{1}{2}\), by Lemma 2.4 (taking \(p = 2\) and \(B := H^{-1}(\mathbb{T}^d)\)) we can obtain
\(m \in C^{\beta-\frac{1}{2}}([0,T]; H^{-1}(\mathbb{T}^d))\).

\(\square\)
3. The coupled time-fractional forward-backward system.

3.1. A formal derivation of the coupled system. In this section we introduce the coupled fractional forward-backward system and the corresponding variational interpretation.

Consider the following optimal control problem

\[
\inf_{v,m} J^{AD} = \inf_{v,m} \int_0^T \int_{\mathbb{T}_d} \frac{1}{2} |v|^2 m + F(x, m) \, dxdt + \int_{\mathbb{T}_d} u_T(x)(I_{(0, t]}^{1-\beta} m)(T)dx,
\]

with \( m \) and \( v \) constrained by the advection-diffusion equation (1) in the weak sense.

We can reformulate this problem as a saddle point problem:

\[
\inf_{v,m} \sup_{u} \mathcal{L}(v, m, u) := \inf_{v,m} \int_0^T \int_{\mathbb{T}_d} \frac{1}{2} |v|^2 m + F(x, m) \, dxdt + \int_{\mathbb{T}_d} u_T(x)(I_{(0, t]}^{1-\beta} m)(T)dx,
\]

\[
+ \sup_{u} \left\{ -\int_0^T \int_{\mathbb{T}_d} u(\partial^\beta_{(0, t]} m - \Delta m + \text{div}(mv)) \, dxdt \right\}.
\]

As in [6], we first show at a formal level the connection between the coupled forward-backward system and the control problem of advection-diffusion equation via Lagrange multiplier methods, assuming that the optimal control problem is convex. Let \((v, m, u)\) solves

\[
\inf_{v,m} \sup_{u} \mathcal{L}(v, m, u),
\]

we claim that \((u, m)\) is a solution to the coupled system (10) with \( v = -Du \). Indeed, via integration by parts, we have:

\[
\inf_{v,m} \sup_{u} \mathcal{L}(v, m, u)
\]

\[
\quad = \inf_{v,m} \int_0^T \int_{\mathbb{T}_d} \frac{1}{2} |v|^2 m + F(x, m) \, dxdt + \int_{\mathbb{T}_d} m(-\partial^\beta_{(0, t]} u + \Delta u + v \nabla u) \, dxdt
\]

\[
\quad + \int_{\mathbb{T}_d} m(0, x)(I_{(0, t]}^{1-\beta} u)(0)dx + \int_{\mathbb{T}_d} (u_T(x) - u(T, x))(I_{(0, t]}^{1-\beta} m)(T)dx.
\]

By convexity of the Lagrangian and Legendre-Fenchel theory,

\[
\inf_{v,m} \sup_{u} \mathcal{L} = \sup_{u} \inf_{v,m} \mathcal{L}.
\]

Given a variation \( \delta v \) with initial condition \( \delta v(0, x) = 0 \), consider first variation,

\[
\frac{\delta \mathcal{L}}{\delta v}(v, u) = \lim_{h \to 0} \frac{\mathcal{L}(v + h\delta v, m, u) - \mathcal{L}(v, m, u)}{h}
\]

\[
= \int_0^T \int_{\mathbb{T}_d} (v + \nabla u) \delta v \cdot m \, dxdt.
\]
We have obtained that optimal control is \( v = -\nabla u \) and we have \( \frac{1}{2}|v|^2 + v \cdot \nabla u = -\frac{1}{2}|\nabla u|^2 \). Then
\[
\inf_{m} \sup_{u} \int_{\mathbb{T}_d} -\frac{1}{2}|\nabla u|^2 m + F(x, m) \, dxdt + \int_{\mathbb{T}_d} m(-\partial^\beta_{[t,T]} u + \Delta u) \, dxdt
+ \int_{\mathbb{T}_d} m(0, x)(I_{[t,T]}^{1-\beta} u)(0)dx + \int_{\mathbb{T}_d} (u_T(x) - u(T, x))(I_{(0,t]}^{1-\beta} m)(T)dx.
\]
Take a variation \( \delta m \) with boundary condition \( \delta m(0, x) = 0 \) and recall that \( \frac{\partial F(x, m)}{\partial m} = f(x, m) \), we can get
\[
\frac{\delta L}{\delta m} = \int_{0}^{T} \int_{\mathbb{T}_d} \delta m(f(x, m) - \frac{1}{2}|\nabla u|^2 + \Delta u - \partial^\beta_{[t,T]} u)dxdt
+ \int_{\mathbb{T}_d} (u_T(x) - u(T, x))(I_{(0,t]}^{1-\beta} \delta m)(T)dx.
\]
We conclude that \((u, m)\) is solution to the HJ equation
\[
\begin{align*}
\partial^\beta_{[t,T]} u - \Delta u + \frac{1}{2}|\nabla u|^2 &= f(x, m), \quad (t, x) \in (0, T) \times \mathbb{T}_d
\end{align*}
\]
\[u(T, x) = u_T(x).\] (19)

### 3.2. Duality of two optimal control problems.
We now show the duality between two optimal control problems constrained, respectively, by a fractional HJ equation and a fractional advection-diffusion equation. We start from introducing a control problem for HJ equation. Denote by \( \mathcal{K}_0 \) the set of maps \( u \in C^2([0, T] \times \mathbb{T}_d) \) such that \( u(T, x) = u_T(x) \) and define, on \( \mathcal{K}_0 \), the functional
\[
\mathcal{A}(u) = \int_{0}^{T} \int_{\mathbb{T}_d} F^*(x, -\partial^\beta_{[t,T]} u - \Delta u + \frac{1}{2}|\nabla u|^2) \, dxdt - \int_{\mathbb{T}_d} m_0(I_{[0,t]}^{1-\beta} u)(0)dx,
\]
the Hamiltonian is defined as \( H(x, p) = \frac{1}{2}|p|^2 \). Here \( F^* \) defines the Legendre-Fenchel transform of \( F(x, m) \) such that \( \forall \alpha: \)
\[
F^*(x, \alpha) = \sup_{m} \{ \alpha m - F(x, m) \}.
\]
This is in fact a control problem of the Hamilton Jacobi equation:
\[
J^HJ(\alpha) = \int_{0}^{T} \int_{\mathbb{T}_d} F^*(x, \alpha(x, m)) \, dxdt - \int_{\mathbb{T}_d} m_0(I_{[0,t]}^{1-\beta} u)(0)dx,
\]
\[J^HJ(\alpha) = \begin{align*}
\partial^\beta_{[t,T]} u - \Delta u + \frac{1}{2}|\nabla u|^2 &= \alpha(x, m),
\end{align*}
\[u(T, x) = u_T(x).\] (21)

Next we reformulate the control problem \( J^{AD} \). To linearize the constraint, we use a standard trick in optimal transport theory and introduce the variable
\[
w(t, x) = v(t, x)m(t, x).
\]
Then the advection-diffusion equation can be written as:
\[
\begin{align*}
\partial^\beta_{[0,t]} m - \Delta m + \text{div}(w) &= 0, \quad (t, x) \in (0, T) \times \mathbb{T}_d, \\
m(0, x) &= m_0(x),
\end{align*}
\]
(22)
where the solution is in the sense of distributions. The Legendre-Fenchel transform of Hamiltonian can be written as:

\[
H^*(x, -\frac{w}{m}) = \begin{cases} 
\frac{1}{2} \left| -\frac{w}{m} \right|^2 & \text{if } m(t, x) > 0, \\
0 & \text{if } (m, w) = (0, 0), \\
+\infty & \text{otherwise.}
\end{cases}
\]

Let us denote \( \mathcal{K}_1 \) the set of pairs \((m, w)\) \(\in L^1((0, T) \times \mathbb{T}^d) \times L^1((0, T) \times \mathbb{T}^d, \mathbb{R}^d)\) such that \(m(t, x) > 0, \int_{\mathbb{T}^d} m(t, x) dx = 1\) for a.e. \(t \in (0, T)\). On the set \( \mathcal{K}_1 \), define the following functional

\[
\mathcal{B}(m, w) = \int_0^T \int_{\mathbb{T}^d} m(t, x) H^*(x, -\frac{w(t, x)}{m(t, x)}) + F(x, m(t, x)) dxdt \\
+ \int_{\mathbb{T}^d} w_T(x)(I_{0, \beta}^1 m(T) dx.
\]

Since \(H^*\) and \(F\) are bounded from below and \(m \geq 0\) a.e. (from Theorem 2.8), the first integral in \(\mathcal{B}(m, w)\) is well defined in \(\mathbb{R} \cup \{+\infty\}\).

We proceed to our main duality result.

**Theorem 3.1.** We have

\[
\inf_{u \in \mathcal{K}_0} \mathcal{A}(u) = -\min_{(m, w) \in \mathcal{K}_1} \mathcal{B}(m, w). 
\] (23)

**Proof.** Let \(E_0 = C^2([0, T] \times \mathbb{T}^d)\) and \(E_1 = C([0, T] \times \mathbb{T}^d, \mathbb{R}) \times C([0, T] \times \mathbb{T}^d, \mathbb{R}^d)\). Define on \(E_0\) the functional

\[
\mathcal{F}(u) = -\int_{\mathbb{T}^d} m_0(x)(I_{0, \beta}^1 u)(0) dx + \chi_S(u),
\]

where \(\chi_S\) is the characteristic function of the set \(S = \{u \in E_0, u(T, \cdot) = u_T\}\), i.e., \(\chi_S(u) = 0\) if \(u \in S\) and \(+\infty\) otherwise. For \((a, b) \in E_1\), we define:

\[
\mathcal{G}(a, b) = \int_0^T \int_{\mathbb{T}^d} F^*(x, -a(t, x) + H(x, b(t, x))) dxdt.
\]

Let \(\Lambda : E_0 \to E_1\) be the bounded linear operator defined by:

\[
\Lambda(u) = (-\partial_{[t, T]}^\beta u + \Delta u, \nabla u).
\]

Then we obtain

\[
\inf_{u \in \mathcal{K}_0} \mathcal{A}(u) = \inf_{u \in \mathcal{K}_0} \{\mathcal{F}(u) + \mathcal{G}(\Lambda(u))\}.
\]

It follows by Fenchel-Rockafellar duality theorem that

\[
\inf_{u \in \mathcal{K}_0} \{\mathcal{F}(u) + \mathcal{G}(\Lambda(u))\} = \max_{(m, w) \in E_1^*} \{-\mathcal{F}^*(\Lambda^*(m, w)) - \mathcal{G}^*(-(m, w))\}
\]

where \(E_1^*\) is the dual space of \(E_1\), i.e., the set of vector valued Radon measures \((m, w)\) over \([0, T] \times \mathbb{T}^d\) with values in \(\mathbb{R} \times \mathbb{R}^d\), \(E_0^*\) is the dual space of \(E_0\), \(\Lambda^* : E_1^* \to E_0^*\) is the dual operator of \(\Lambda\) and \(\mathcal{F}^*\) and \(\mathcal{G}^*\) are the convex conjugate of \(\mathcal{F}\) and \(\mathcal{G}\).
respective.

\[ \mathcal{F}^*(\Lambda^*(m, w)) \]
\[ = \sup_u \{(\Lambda^*(m, w), u) - \mathcal{F}(u)\} \]
\[ = \sup_u \{(m, w), \Lambda u - \mathcal{F}(u)\} \]
\[ = \sup_u \int_0^T \int_{\mathbb{T}^d} m(-\partial_{(0,1)}^\beta u + \Delta u)dxdt + \int_0^T \int_{\mathbb{T}^d} u\nabla u dxdt - \mathcal{F}(u) \]
\[ = \sup_u \int_0^T \int_{\mathbb{T}^d} u(-\partial_{(0,1)}^\beta m + \Delta m - \nabla w)dxdt + \int_{\mathbb{T}^d} (m_0(x) - m(0, x))(I_{1-\beta}^t u)(0)dx \]
\[ - \chi_S(u) + \int_{\mathbb{T}^d} u(T, x)(I_{(0,\beta)}^1 m)(T)dx \].

Therefore

\[ \mathcal{F}^*(\Lambda^*(m, w)) = \begin{cases} \int_{\mathbb{T}^d} u_T(x)(I_{(0,\beta)}^{1-\beta} m)(T)dx & \text{if } m, w \text{ satisfy equation (22),} \\ +\infty & \text{otherwise.} \end{cases} \]

Moreover, for any \((m, w) \in \mathbb{R} \times \mathbb{R}^d\),

\[ \mathcal{G}^*(m, w) \]
\[ = \sup_{a \in \mathbb{R}, b \in \mathbb{R}^d} \left\{ \int_0^T \int_{\mathbb{T}^d} (am + \langle b, w \rangle)dxdt - \mathcal{G}(a, b) \right\} \]
\[ = \sup_{a \in \mathbb{R}, b \in \mathbb{R}^d} \int_0^T \int_{\mathbb{T}^d} (am + \langle b, w \rangle)dxdt - \int_0^T \int_{\mathbb{T}^d} F^*(x, -a + H(x, b))dxdt \]
\[ = \sup_{a \in \mathbb{R}, b \in \mathbb{R}^d} \int_0^T \int_{\mathbb{T}^d} mH(x, b) - am + \langle b, w \rangle - F^*(x, a)dxdt. \]

Hence, for \((-m) > 0\), due to the convexity of \(F\), we have

\[ \mathcal{G}^*(m, w) \]
\[ = \sup_{a \in \mathbb{R}, b \in \mathbb{R}^d} \int_0^T \int_{\mathbb{T}^d} (-m) \left( \langle b, -\frac{w}{m} \rangle - H(x, b) \right) + (-m)a - F^*(x, a)dxdt \]
\[ = \int_0^T \int_{\mathbb{T}^d} (-m)H^*(x, -\frac{w}{m}) + F(x, -m)dxdt. \]

Therefore, for \(m > 0\),

\[ \max_{(m, w) \in E_t^1} \{-\mathcal{F}^*(\Lambda^*(m, w)) - \mathcal{G}^*(-(m, w))\} \]
\[ = \max_{(m, w) \in E_t^1} \left\{ \int_0^T \int_{\mathbb{T}^d} -mH^*(x, -\frac{w}{m}) - F(x, m)dxdt - \int_0^T u_T(x)(I_{(0,\beta)}^{1-\beta} m)(T)dx \right\} \]
\[ = -\min_{(m, w) \in E_t^1} \left\{ \int_0^T \int_{\mathbb{T}^d} mH^*(x, -\frac{w}{m}) + F(x, m)dxdt + \int_0^T u_T(x)(I_{(0,\beta)}^{1-\beta} m)(T)dx \right\} \]
\[ = -\min_{(m, w) \in E_t^1} \mathcal{B}(m, w). \]

The minimum is taken over the \(L^1\) maps \((m, w)\) such that \(m(t, x) \geq 0\) a.e. and equation (22) is satisfied in the sense of distributions. Since \(\int_{\mathbb{T}^d} m_0 dx = 1\), by Theorem 2.8 we have \(\int_{\mathbb{T}^d} m(t, x) dx = 1\) for all \(t \in [0, T]\). Hence the pair
\[(m, w)\] belongs to the set \(K_1\) and we have proved the duality of the optimal control problems.

\[\square\]

3.3. **Verification arguments.** In the next theorem, we show the connection between the optimal control problems for the fractional advection-diffusion and HJ equations and the coupled system (2).

**Theorem 3.2.** Assume that \((\hat{u}, \hat{m})\) is of class \(C^1([0, T] \times \mathbb{T}^d) \times \mathbb{R}^d\), with \(\hat{m}(0, x) = m_0\) and \(\hat{u}(T, x) = u_T(x)\). Then the following statements are equivalent:

(i) \((\hat{u}, \hat{m})\) is a solution to the fractional coupled system (2).

(ii) The solution \(\hat{u}(x, t)\) is optimal for \(\inf_u A\).

(iii) The control \(\hat{v} = -\nabla \hat{u}(t, x)\) and \(\hat{m}\) are optimal for \(\min_{(m, w)} B\) where \(\hat{w} = \hat{v} \hat{m}\), \(\hat{m}\) is the solution of advection-diffusion equation (1).

**Proof.** The proof is by verification arguments.

"i \(\Rightarrow ii"\): assume that \((\hat{m}, \hat{u})\) is solution to coupled system (2). Denote by \(\hat{\alpha}\) and \(\alpha\) respectively

\[
\hat{\alpha} = \partial_{[t, T]}^{\beta} \hat{u} - \Delta \hat{u} + \frac{1}{2}|\nabla \hat{u}|^2,
\]

\[
\alpha = \partial_{[t, T]}^{\beta} u - \Delta u + \frac{1}{2}|\nabla u|^2,
\]

\[
\hat{u}(T, x) = u(T, x) = u_T(x).
\]

Thus we can write

\[
A(u) = J_{\beta}^{HJ}(\alpha) = \int_0^T \int_{\mathbb{T}^d} F^*(x, \alpha) \, dx \, dt - \int_{\mathbb{T}^d} m_0(x) (I^{1-\beta}_{[t, T]} u)(0) \, dx.
\]

By definition of Legendre transform \(F^*(x, \alpha)\) we know it is convex with regard to \(\alpha\), hence

\[
J_{\beta}^{HJ}(\alpha) \geq J_{\beta}^{HJ}(\hat{\alpha}) + \int_0^T \int_{\mathbb{T}^d} \partial_\alpha F^*(x, \hat{\alpha})(\alpha - \hat{\alpha}) \, dx \, dt
\]

\[
- \int_{\mathbb{T}^d} ((I^{1-\beta}_{[t, T]} u)(0) - (I^{1-\beta}_{[t, T]} \hat{u})(0)) \, dm_0(x)
\]

\[
= J_{\beta}^{HJ}(\hat{\alpha}) + \int_0^T \int_{\mathbb{T}^d} \partial_\alpha F^*(x, \hat{\alpha})(\alpha - \hat{\alpha}) \, dx \, dt
\]

\[
- \int_{\mathbb{T}^d} ((I^{1-\beta}_{[t, T]} u)(0) - (I^{1-\beta}_{[t, T]} \hat{u})(0)) \, dm_0(x) \, dx.
\]

Since, by definition, \(\partial_\alpha F^*(x, \hat{\alpha}) = \hat{m}\) and \(|\nabla u|^2 - |\nabla \hat{u}|^2 \geq 2\nabla \hat{u}(u - \hat{u})\), we obtain

\[
J_{\beta}^{HJ}(\alpha)
\]

\[
\geq J_{\beta}^{HJ}(\hat{\alpha}) + \int_0^T \int_{\mathbb{T}^d} \hat{m}(\partial_{[t, T]}^{\beta} (u - \hat{u})) + \Delta (u - \hat{u}) + \nabla \hat{u}(u - \hat{u}) \, dx \, dt
\]

\[
- \int_{\mathbb{T}^d} ((I^{1-\beta}_{[t, T]} u)(0) - (I^{1-\beta}_{[t, T]} \hat{u})(0)) \, dm_0(x) \, dx
\]

\[
= J_{\beta}^{HJ}(\hat{\alpha}) + \int_0^T \int_{\mathbb{T}^d} (u - \hat{u})(\partial_{[t, T]}^{\beta} \hat{m} - \Delta \hat{m} + \nabla \hat{m} \nabla \hat{u}) \, dx \, dt
\]

\[
- \int_{\mathbb{T}^d} (u(T, x) - \hat{u}(T, x))(I^{1-\beta}_{[t, T]} \hat{u})(0) \, dx
\]

\[
+ \int_{\mathbb{T}^d} ((I^{1-\beta}_{[t, T]} \hat{u} - I^{1-\beta}_{[t, T]} u)(0))(\hat{m}(0, x) - m_0(x)) \, dx \, dt.
\]
As \((\hat{m}, \hat{u})\) is solution to coupled system (2) then
\[
\partial^\beta_{(0,T)} \hat{m} - \Delta \hat{m} + \text{div}(\hat{m} \nabla \hat{u}) = 0, \quad \hat{m}(0, x) = m_0,
\]
is satisfied in the sense of distributions. Hence we get \(J^{HJ}(\alpha) \geq J^{HJ}(\hat{\alpha})\). From the uniqueness of solution to the fractional HJ equation we conclude: \(A(u) \geq A(\hat{u})\).

\(\text{“ii} \Rightarrow \text{“i”} \): Define \(\hat{m}(t, x) = \partial_\alpha F^*(x, \hat{\alpha}(t, x))\), then \(\hat{\alpha}(t, x) = f(x, \hat{m}(t, x))\). Take a smooth function \(\delta \alpha \in C^1([0, T] \times \mathbb{T}^d)\), denote by \(u_h\) the solution to the equation
\[
\begin{align*}
\partial^\beta_{(t,T)} u - \Delta u + \frac{1}{2} |\nabla u|^2 &= \hat{\alpha} + h \delta \alpha, \quad (t, x) \in (0, T) \times \mathbb{T}^d \\
u(T, x) &= u_T(x).
\end{align*}
\]
We remind that \(\hat{u}\) is the solution to the system
\[
\begin{align*}
\partial^\beta_{(t,T)} u - \Delta u + \frac{1}{2} |\nabla u|^2 &= \hat{\alpha}, \quad (t, x) \in (0, T) \times \mathbb{T}^d \\
u(T, x) &= u_T(x).
\end{align*}
\]
Thus \((u_h - \hat{u})/h\) converges to a smooth function \(\phi\) which is the solution to the linearized system:
\[
\begin{align*}
\partial^\beta_{(t,T)} \phi - \Delta \phi + \nabla \hat{u} \cdot \nabla \phi &= \delta \alpha(t, x), \quad (t, x) \in (0, T) \times \mathbb{T}^d \\
\phi(T, x) &= 0.
\end{align*}
\]
Calculate the first variation we can obtain
\[
\frac{\delta J^{HJ}(\hat{\alpha})}{\delta \alpha} = \lim_{h \to 0} \frac{J^{HJ}(\hat{\alpha} + h \delta \alpha) - J^{HJ}(\hat{\alpha})}{h}
\]
\[
= \int_0^T \int_{\mathbb{T}^d} \lim_{h \to 0} \frac{F^*(x, \hat{\alpha} + h \delta \alpha) - F^*(x, \hat{\alpha})}{h} dx dt - \int_{\mathbb{T}^d} (I_{(t,T)}^{1-\beta}) \phi(0)m_0(x) dx
\]
\[
= \int_0^T \int_{\mathbb{T}^d} \hat{m} \cdot \delta \alpha dx dt - \int_{\mathbb{T}^d} (I_{(t,T)}^{1-\beta}) \phi(0)m_0(x) dx
\]
\[
= \int_0^T \int_{\mathbb{T}^d} \hat{m}(\partial^\beta_{(t,T)} \phi - \Delta \phi + \nabla \hat{u} \cdot \nabla \phi) dx dt - \int_{\mathbb{T}^d} (I_{(t,T)}^{1-\beta}) \phi(0)m_0(x) dx
\]
\[
= \int_0^T \int_{\mathbb{T}^d} \phi(\partial^\beta_{(t,T)} \hat{m} - \Delta \hat{m} + \text{div}(\hat{m} \nabla \hat{u})) dx dt - \int_{\mathbb{T}^d} (I_{(t,T)}^{1-\beta}) \phi(0)(m_0(x) - \hat{m}(0, x)) dx.
\]
Since by density argument this holds for any test function \(\phi \in C^3\) with \(\phi(T, x) = 0\), thus we obtain \(\hat{m}\) is a weak solution to the advection-diffusion equation in the coupled system (2) with \(\hat{m}(0, x) = m_0(x)\).

\(\text{“(i)”} \Rightarrow \text{“(iii)”}\) This can be proven by verification argument, completely analogous to the reasoning above, hence we omit it. \(\Box\)

3.4. Existence and uniqueness of solution.

**Theorem 3.3. (Existence)** Suppose \(f(x, \cdot)\) is a regularizing coupling such that it maps the subset \(X\) of \(L^2([0, T]; L^2(\mathbb{T}^d))\) into \(W^{1,\infty}([0, T]; W^{1,\infty}(\mathbb{T}^d))\). \(f\) is continuous from \(X\) into \(C([0, T]; C(\mathbb{T}^d))\). Then the system (2) has at least one solution \((u, m) \in C^1([0, T]; C^2(\mathbb{T}^d)) \times \Xi\).

**Proof.** We obtain the existence of solution via Schauder fixed point theorem.

**Boundedness:** Let \(C\) be a large constant to be chosen below and denote by \(C\) the
set of maps \(\mu \in L^2([0, T]; L^2(\mathbb{T}^d))\) such that
\[
\|\mu\|_{L^2([0, T]; L^2(\mathbb{T}^d))} \leq C,
\]
constant \(C\) is to be fixed later. Then \(C\) is a convex closed set set of \(L^2([0, T]; L^2(\mathbb{T}^d))\).

To any \(\mu \in C\) we associate \(m = \Psi(\mu)\). Let \(u\) be the solution to
\[
\begin{cases}
\partial^\beta_{(t, \mathbb{T}^d)} u - \Delta u + \frac{1}{2} |\nabla u|^2 = f(x, \mu), & (t, x) \in (0, T) \times \mathbb{T}^d \\
u(T, x) = u_T(x),
\end{cases}
\]
and \(\|\mu\| \leq c_1\). It has been proven that in this case there is a unique classical solution \(u\). We denote with constant \(c_3 > 0\) such that \(\forall t \in [0, T], \|\nabla u(t, \cdot)\| \leq c_3\). Then \(m\)

solves the advection-diffusion equation
\[
\begin{cases}
\partial^\beta_{(0, \mathbb{T}^d)} u - \Delta m + \text{div}(-m \nabla u) = 0, & (t, x) \in (0, T) \times \mathbb{T}^d \\
m(0, x) = m_0(x),
\end{cases}
\]
By a priori estimates we have \(m \in \Xi\) and the bound
\[
\|m\|_{L^2([0, T]; L^2(\mathbb{T}^d))} \leq c_2(c_3 T)^{\frac{1}{2}} \left(\sum_{k=0}^\infty \left(\frac{c_3^2 T^\beta}{2}\right)^k \frac{1}{\Gamma(1 + \beta k)}\right)^{\frac{1}{2}} \|m_0(x)\|_2.
\]

Therefore we can choose
\[
C := \max\{c_1, c_2(c_3 T)^{\frac{1}{2}} \left(\sum_{k=0}^\infty \left(\frac{c_3^2 T^\beta}{2}\right)^k \frac{1}{\Gamma(1 + \beta k)}\right)^{\frac{1}{2}} \|m_0(x)\|_2\},
\]
such that we have \(\Psi : \mu \mapsto m\) maps from \(C\) into itself.

COMPACTNESS: Consider the triplet (Gelfand triple):
\[
H^1(\mathbb{T}^d) \hookrightarrow L^2(\mathbb{T}^d) \hookrightarrow H^{-1}(\mathbb{T}^d),
\]
\(H^1(\mathbb{T}^d)\) is compactly embedded in \(L^2(\mathbb{T}^d)\), by using the generalized Aubin-Lions argument (Theorem 4.1 in [30]) we obtain that:
\[
\Xi \text{ is compactly embedded in } L^2([0, T]; L^2(\mathbb{T}^d)).
\]

CONTINUITY: Consider a sequence \(\mu_n \in C\) and denote by \((u_n, m_n)\) and \((u, m)\)
the solution obtained via \(\mu_n\) and \(\mu\), respectively. It has been shown by Kolokolstov et al. ([26], Theorem 11) that there exist uniform bound \(c_4\) such that
\[
\sup_{t \in [0, T]} \|u_n\|_{C^2(\mathbb{T}^d)} \leq c_4.
\]

Hence we conclude that \(u_n \rightarrow u\) and \(\nabla u_n\) converges locally uniformly to \(u\) and \(\nabla u\), respectively. Moreover, from the estimates in the proof of Theorem 2.8 we have again a uniform bound for \(m_n\):
\[
I^{1-\beta}_{(0, T)}(\|m_n\|_2^2) + \int_0^T \int_{\mathbb{T}^d} |\nabla m_n|^2 \, dx \, dt \leq c_4 T \left(\sum_{k=0}^\infty \left(\frac{c_3^2 T^\beta}{2}\right)^k \frac{1}{\Gamma(1 + \beta k)}\right) \|m_0(x)\|_2^2.
\]
We can then extract a weakly convergent subsequence of \(m_n\), the limit will be a weak solution of the advection-diffusion equation in (2). Since we have proved the weak solution is unique in \(\Xi\) for fixed \(\nabla u\), thus \(m\) is the limit and the mapping \(\Psi\) is continuous.

As \(\Psi\) maps a Banach space continuously into a convex compact subset, we conclude with Schauder fixed point theorem it admits at least one fixed point \(\Psi(m) = m\). It follows that the pair \((u, m)\) is a solution to the system. \(\square\)
Theorem 3.4. (Uniqueness) Under assumptions (H1)-(H3), the solution constructed in theorem 3.3 for the coupled system (2) is unique.

Proof. Suppose \((u_1, m_1)\) and \((u_2, m_2)\) are two solutions of the system (2), denote \(\bar{u} = u_1 - u_2\) and \(\bar{m} = m_1 - m_2\), we obtain

\[
- \partial^\beta_{(t,T)} \bar{u} + \Delta \bar{u} + \frac{1}{2} (|\nabla u_2|^2 - |\nabla u_1|^2) - (f(x, m_2) - f(x, m_1)) = 0,
\]

and

\[
\partial^\beta_{(0,t)} \bar{m} - \Delta \bar{m} - \text{div}(m_1 \nabla u_1 - m_2 \nabla u_2) = 0.
\]

By definition of \(\bar{u}\) and \(\bar{m}\) we have \(\bar{u}(T, x) = 0\) and \(\bar{m}(0, x) = 0\), therefore

\[
\int_0^T \int_{T_4} (-\partial^\beta_{(t,T)} \bar{u}) \bar{m} dx dt + \int_0^T \int_{T_4} (\partial^\beta_{(0,t)} \bar{m}) \bar{u} dx dt = \int_{T_4} (\bar{u}(T, x)(I^{1-\beta}_{0,T} \bar{m})(T) - \bar{m}(0, x)(I^{1-\beta}_{t,T} \bar{u})(0)) dx = 0.
\]

Multiplying both sides of (29) by \(\bar{m}\) and (30) by \(\bar{u}\), add up and then integrate in the domain \((0, T) \times T^d\) we obtain

\[
0 = \int_0^T \int_{T_4} \frac{1}{2} \bar{m}(|\nabla u_1|^2 - |\nabla u_2|^2) - \bar{m}(f(x, m_1) - f(x, m_2)) dx dt + \int_0^T \int_{T_4} \bar{u} \text{div}(m_1 \nabla u_1 - m_2 \nabla u_2) dx dt
\]

\[
= \int_0^T \int_{T_4} \frac{1}{2} \bar{m}(|\nabla u_1|^2 - |\nabla u_2|^2) - \bar{m}(f(x, m_1) - f(x, m_2)) dx dt - \int_0^T \int_{T_4} \bar{u} \cdot (m_1 \nabla u_1 - m_2 \nabla u_2) dx dt,
\]

and also note that

\[
\frac{1}{2} \bar{m}(|\nabla u_1|^2 - |\nabla u_2|^2) - \langle \nabla \bar{u}, m_1 \nabla u_1 - m_2 \nabla u_2 \rangle = - \frac{1}{2} \bar{m} |\nabla u_1 - \nabla u_2|^2 \leq 0,
\]

so that

\[
\int_0^T \int_{T_4} \bar{m}(f(x, m_1) - f(x, m_2)) dx dt \leq 0.
\]

Comparing with assumption (H2) we obtain \(\bar{m} = 0\) and by uniqueness of solution to the Hamilton-Jacobi equation we obtain \(\bar{u} = 0\) and the solution to the fractional coupled system (2) is unique.

4. Discussions. In this paper we studied an optimal control problem of time-fractional advection-diffusion equation which leads to a coupled system with a backward time-fractional Hamilton Jacobi equation as the adjoint equation. Using methods from convex analysis we have shown the equivalence between some optimization problems and the coupled PDE system. We obtained existence and uniqueness results under quite restrictive assumptions. It should be natural to consider the system with local coupling \(f(x, m)\) (without regularization).

Another important direction would be to explore further the relationship between the coupled system (2) and the time-fractional mean field game system (11) and
try to find some entropy minimization and gradient flow formulation for both. Recently Cirant and Goffi [16] studied a nonlocal (in space) MFG system:

\[
\begin{align*}
-\partial_t u + (-\Delta)^s u + \frac{1}{2} |\nabla u|^2 &= f(x, m), \quad (t, x) \in (0, T) \times \mathbb{T}^d \\
\partial_t m + (-\Delta)^s m + \text{div}(-m \nabla u) &= 0,
\end{align*}
\]

(31)

where \((-\Delta)^s\) denotes the fractional Laplacian of order \(s\), \(0 < s < 1\). A stationary nonlocal in space MFG system has been studied by Cesaroni et al. in [15]. The methods in our paper can be extended to the generalized system:

\[
\begin{align*}
\partial_t^\alpha u + (-\Delta)^s u + \frac{1}{2} |\nabla u|^2 &= f(x, m), \quad (t, x) \in (0, T) \times \mathbb{T}^d \\
\partial_t^\beta m + (-\Delta)^s m + \text{div}(-m \nabla u) &= 0, \\
m(0, x) = m_0(x), \quad u(T, x) = u_T(x),
\end{align*}
\]

(32)

Also, it would be of great interest to construct a mean field game system with both space and time nonlocal structure.

REFERENCES

[1] Y. Achdou, M. Bardi and M. Cirant, Mean field games models of segregation, *Math. Models Methods Appl. Sci.*, 27 (2017), 75–113.

[2] M. Allen, L. Caffarelli and A. Vasseur, A parabolic problem with a fractional time derivative, *Arch. Ration. Mech. Anal.*, 221 (2016), 603–630.

[3] M. Annunziato, A. Borzì, M. Magdziarz and A. Weron, A fractional Fokker-Planck control framework for subdiffusion processes, *Optimal Control Appl. Methods*, 37 (2016), 290–304.

[4] H. Antil, E. Otárola and A. J. Salgado, A space-time fractional optimal control problem: Analysis and discretization, *SIAM J. Control Optim.*, 54 (2016), 1295–1328.

[5] J.-D. Benamou and Y. Brenier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem, *Numer. Math.*, 84 (2000), 375–393.

[6] J.-D. Benamou, G. Carlier and F. Santambrogio, Variational mean field games, in *Active Particles, Model. Simul. Sci. Eng. Technol.*, Birkhäuser/Springer, Cham, 1 (2017), 141–171.

[7] J.-D. Benamou, G. Carlier, S. D. Marino and L. Nenna, An entropy minimization approach to second-order variational mean-field games, *Math. Models Methods Appl. Sci.*, 29 (2019), 1553–1583.

[8] J.-P. Bouchaud and A. Georges, Anomalous diffusion in disordered media: Statistical mechanisms, models and physical applications, *Phys. Rep.*, 195 (1990), 127–293.

[9] M. Burger, M. D. Francesco, P. A. Markowich and M.-T. Wolfram, Mean field games with nonlinear mobilities in pedestrian dynamics, *Discrete Contin. Dyn. Syst. Ser. B*, 19 (2014), 1311–1333.

[10] F. Camilli, R. D. Maio and E. Iacomini, A Hopf-Lax formula for Hamilton-Jacobi equations with Caputo time derivative, *J. Math. Anal. Appl.*, 477 (2019), 1019–1032.

[11] F. Camilli and R. D. Maio, A time-fractional mean field game, *Adv. Differential Equations*, 24 (2019), 531–554.

[12] P. Cardaliaguet, *Notes on Mean Field Games from P.L. Lions’ lectures at Collège de France*, 2015.

[13] P. Cardaliaguet, Weak solutions for first order mean field games with local coupling, in *Analysis and Geometry in Control Theory and its Applications*, Springer INdAM Ser., Springer, Cham, 11 (2015), 111–158.

[14] P. Cardaliaguet, P. J. Graber, A. Porretta and D. Tonon, Second order mean field games with degenerate diffusion and local coupling, *NoDEA Nonlinear Differential Equations Appl.*, 22 (2015), 1287–1317.

[15] A. Cesaroni, M. Cirant, S. Dipierro, M. Novaga and E. Valdinoci, On stationary fractional mean field games, *J. Math. Pures Appl.*, 122 (2019), 1–22.

[16] M. Cirant and A. Goffi, On the existence and uniqueness of solutions to time-dependent fractional MFG, *SIAM J. Math. Anal.*, 51 (2019), 913–954.

[17] D. A. Gomes and J. Saúde, Mean field games models—A brief survey, *Dyn. Games Appl.*, 4 (2014), 110–154.
[18] Y. Giga and T. Namba, Well-posedness of Hamilton-Jacobi equations with Caputo’s time fractional derivative, *Comm. Partial Differential Equations*, 42 (2017), 1088–1120.
[19] O. Guéant, J.-M. Lasry and P.-L. Lions, Mean field games and applications, in *Paris-Princeton Lectures on Mathematical Finance 2010*, Lecture Notes in Math., 2003, Springer, Berlin, 2011, 205–266.
[20] M. Hahn, K. Kobayashi and S. Umarov, SDEs driven by a time-changed Lévy process and their associated time-fractional order pseudo-differential equations, *J. Theoret. Probab.*, 25 (2012), 262–279.
[21] G. H. Hardy and J. E. Littlewood, Some properties of fractional integrals. I., *Math. Z.*, 27 (1928), 556–606.
[22] M. E. Hernández-Hernández and V. N. Kolokoltsov, Probabilistic solutions to nonlinear fractional differential equations of generalized Caputo and Riemann–Liouville type, *Stochastics*, 90 (2018), 224–255.
[23] M. Huang, P. E. Caines and R. P. Malhame, Large-population cost-coupled LQG problems with non uniform agents: Individual-mass behaviour and decentralized ε-Nash equilibria, *IEEE Trans. Automat. Control*, 52 (2007), 1560–1571.
[24] J. Kemppainen and R. Zacher, Long-time behaviour of non-local in time Fokker-Planck equations via the entropy method, *Math. Models Methods Appl. Sci.*, 29 (2019), 209–235.
[25] V. Kolokoltsov and M. Veretennikova, A fractional Hamilton-Jacobi Bellman equation for scaled limits of controlled continuous time random walks, *Commun. Appl. Ind. Math.*, 6 (2014), e-484, 18 pp.
[26] V. Kolokoltsov and M. Veretennikova, Well-posedness and regularity of the Cauchy problem for nonlinear fractional in time and space equations, *Fract. Differ. Calc.*, 4 (2014), 1–30.
[27] A. Kubica and M. Yamamoto, Initial-boundary value problems for fractional diffusion equations with time-dependent coefficients, *Fract. Calc. Appl. Anal.*, 21 (2018), 276–311.
[28] J.-M. Lasry and P.-L. Lions, Jeux à champ moyen. II – Horizon fini et contrôle optimal, *C. R. Math. Acad. Sci. Paris*, 343 (2006), 679–684.
[29] J.-M. Lasry and P.-L. Lions, Mean field games, *Jpn. J. Math.*, 2 (2007), 229–260.
[30] L. Li and J. Liu, Some compactness criteria for weak solutions of time fractional PDEs, *SIAM J. Math. Anal.*, 50 (2018), 3963–3995.
[31] L. Li, J.-G. Liu and L. Wang, Cauchy problems for Keller–Segel type time–space fractional diffusion equation, *J. Differential Equations*, 265 (2018), 1044–1096.
[32] A. R. Mészáros and F. J. Silva, A variational approach to second order mean field games with density constraints: The stationary case, *J. Math. Pures Appl.*, 104 (2015), 1135–1159.
[33] R. Metzler and J. Klafter, The random walk’s guide to anomalous diffusion: A fractional dynamics approach, *Phys. Rep.*, 339 (2000), 1–77.
[34] S. Samko, A. Kilbas and O. Marichev, *Fractional Integrals and Derivatives*, Gordon and Breach Science Publishers, Yverdon, 1993.
[35] Q. Tang and F. Camilli, Variational time-fractional mean field games, *Dynamic Games and Applications*, Springer US, 2019, 1–16.
[36] E. Topp and M. Yangari, Existence and uniqueness for parabolic problems with Caputo time derivative, *J. Differential Equations*, 262 (2017), 6018–6046.
[37] R. Zacher, Boundedness of weak solutions to evolutionary partial integro-differential equations with discontinuous coefficients, *J. Math. Anal. Appl.*, 348 (2008), 137–149.
[38] R. Zacher, A De Giorgi-Nash type theorem for time fractional diffusion equations, *Math. Ann.*, 356 (2013), 99–146.

Received January 2019; revised April 2019.

E-mail address: tangqingthomas@gmail.com