

# An operator splitting scheme for the fractional kinetic Fokker-Planck equation

Duong, Manh Hong; Lu, Yulong

DOI:  
[10.3934/dcds.2019250](https://doi.org/10.3934/dcds.2019250)

License:  
None: All rights reserved

Document Version  
Peer reviewed version

Citation for published version (Harvard):  
Duong, MH & Lu, Y 2019, 'An operator splitting scheme for the fractional kinetic Fokker-Planck equation', *Discrete and Continuous Dynamical Systems - Series A*, vol. 39, no. 10, pp. 5707-5727.  
<https://doi.org/10.3934/dcds.2019250>

[Link to publication on Research at Birmingham portal](#)

**Publisher Rights Statement:**  
Checked for eligibility: 09/05/2019

This is a pre-copy-editing, author-produced PDF of an article accepted for publication in *Discrete & Continuous Dynamical Systems - A*, following peer review. The definitive publisher-authenticated version Manh Hong Duong, Yulong Lu. An operator splitting scheme for the fractional kinetic Fokker-Planck equation. *Discrete & Continuous Dynamical Systems - A*, 2019, 39 (10) : 5707-5727, is available online at: <https://dx.doi.org/10.3934/dcds.2019250>

## General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

## Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.

# AN OPERATOR SPLITTING SCHEME FOR THE FRACTIONAL KINETIC FOKKER-PLANCK EQUATION

MANH HONG DUONG AND YULONG LU

**ABSTRACT.** In this paper, we develop an operator splitting scheme for the fractional kinetic Fokker-Planck equation (FKFPE). The scheme consists of two phases: a fractional diffusion phase and a kinetic transport phase. The first phase is solved exactly using the convolution operator while the second one is solved approximately using a variational scheme that minimizes an energy functional with respect to a certain Kantorovich optimal transport cost functional. We prove the convergence of the scheme to a weak solution to FKFPE. As a by-product of our analysis, we also establish a variational formulation for a kinetic transport equation that is relevant in the second phase. Finally, we discuss some extensions of our analysis to more complex systems.

## 1. INTRODUCTION

In this paper, we study the existence of solutions to the following fractional kinetic Fokker-Planck equation (FKFPE)

$$(1.1) \quad \begin{cases} \partial_t f + v \cdot \nabla_x f = \operatorname{div}_v(\nabla \Psi(v)f) - (-\Delta_v)^s f & \text{in } \mathbf{R}^d \times \mathbf{R}^d \times (0, \infty), \\ f(x, v, 0) = f_0(x, v) & \text{in } \mathbf{R}^d \times \mathbf{R}^d, \end{cases}$$

with  $s \in (0, 1]$ . In the above,  $\operatorname{div}$  denotes the divergence operator; the differential operators  $\nabla, \operatorname{div}$  and  $\Delta$  with subscripts  $x$  and  $v$  indicate that these operators act only on the corresponding variables; the operator  $-(-\Delta_v)^s$  is the fractional Laplacian operator on the variable  $v$ , where the fractional Laplacian  $-(-\Delta)^s$ , is defined by

$$-(-\Delta)^s f(x) := -\mathcal{F}^{-1}(|\xi|^{2s} \mathcal{F}[f](\xi))(x).$$

Here  $\mathcal{F}$  denotes the Fourier transform on  $\mathbf{R}^d$ , i.e.  $\mathcal{F}[f](\xi) = \frac{1}{(2\pi)^{d/2}} \int_{\mathbf{R}^d} f(x) e^{-ix \cdot \xi} dx$ . Note that the fractional Laplacian operator with  $0 < s < 1$  is a non-local operator since it can also be expressed as the singular integral

$$-(-\Delta)^s f(x) = -C_{d,s} \int_{\mathbf{R}^d} \frac{f(x) - f(y)}{|x - y|^{d+2s}} dy,$$

where the normalisation constant is given by  $C_{d,s} = s2^{2s}\Gamma(\frac{d+2s}{2})/(\pi^{\frac{d}{2}}\Gamma(1-s))$  and  $\Gamma(t)$  is the Gamma function. See [33] for more equivalent definitions of fractional Laplacian operator.

The equation (1.1) is interesting to us because it can be viewed as the Fokker-Planck (forward Kolmogorov) equation of the following generalized Langevin equation

$$(1.2) \quad \begin{aligned} \frac{dX_t}{dt} &= V_t, \\ \frac{dV_t}{dt} &= -\nabla \Psi(V_t) + dL_t^s, \end{aligned}$$

---

2010 *Mathematics Subject Classification.* Primary: 49S05, 35Q84; Secondary: 49J40.

*Key words and phrases.* Operator splitting methods, variational methods, fractional kinetic Fokker-Planck equation, kinetic transport equation, optimal transportation.

where  $L_t^s$  is the Lévy stable process with exponent  $2s$ . The stochastic differential equation (SDE) (1.2) describes the motion of a particle moving under the influence of a (generalized) frictional force and a stochastic noise and in the absence of an external force field. FKFPE (1.1) is the evolution of the probability distribution of  $(X_t, V_t)$ . In particular, the fractional operator  $-(-\Delta)^s$  is the generator of the process  $L_t^s$ . When  $s = 1$  and  $\Psi(v) = \frac{|v|^2}{2}$ , equation (1.1) becomes the classical kinetic Fokker-Planck (or Kramers) equation (without external force field) which is a local PDE and has been used widely in chemistry as a simplified model for chemical reactions [32, 26] and in statistical mechanics [35, 38]. The non-local Lévy process plays an important role in modelling systems that include jumps and long-distance interactions such as anomalous diffusion or transport in confined plasma [5]. Singular limits of Equation (1.1) with  $\Psi(v) = \frac{|v|^2}{2}$  was studied in [12], see also [11] for a similar result for the same equation but on a spatially bounded domain. In a recent work [1], the authors have extended [12] to a system that contains an additional external force field and they have also proved its well-posedness by the means of the Lax-Milgram theorem. We will prove the existence of solutions of (1.1) for a general  $\Psi$  based on the trick of operator splitting. For more recent developments on PDEs involving the fractional Laplacian operator, we refer the interested reader to expository surveys [43, 42, 41].

The aim of this paper is to develop a variational formulation for approximating solutions to equation (1.1). The theory of variational formulation for PDEs took off with the introduction of Wasserstein gradient flows by the seminal work of Jordan-Kinderlehrer-Otto [30]. Such a variational structure has important applications for the analysis of an evolution equation such as providing general methods for proving well-posedness [4] and characterizing large time behaviour (e.g., [10]), giving rise to natural numerical discretizations (e.g., [22]), and offering techniques for the analysis of singular limits (e.g., [39, 40, 6, 18]). There are now a significantly large number of papers in exploring variational structures for local PDEs, see the aforementioned papers and references therein as well as the monographs [4, 44] for more details. However, variational formulations for non-local PDEs are less understood. Erbar [24] showed that the fractional heat equation is a gradient flow of the Boltzmann entropy with respect to a new modified Wasserstein distance that is built from the Lévy measure and based on the Benamou-Brenier variant of the Wasserstein distance. Bowles and Agueh [8] proved the existence of the fractional Fokker-Planck equation

$$(1.3) \quad \begin{cases} \partial_t f = \operatorname{div}_v(\nabla \Psi(v)f) - (-\Delta_v)^s f & \text{in } \mathbf{R}^d \times (0, \infty), \\ f(v, 0) = f_0(v) & \text{in } \mathbf{R}^d, \end{cases}$$

which can be viewed as the spatially homogeneous version of equation (1.1) or the fractional heat equation with a drift. Erbar's proof is variational based on the so-called "evolution variational inequality" concept introduced in [4]. However, it seems that his method can not be extended to the fractional Fokker-Planck equation since the distance that he introduced was particularly tailored for the Boltzmann entropy. Instead, Bowles and Agueh's proof is "semi-variational" based on a novel splitting argument which we sketch now. They split up the original dynamics (1.3) into two processes: a fractional diffusion process, namely  $\partial_t f = -(-\Delta)^s f$ , and a transport process in the field of the potential  $\Psi$ , namely  $\partial_t f = \operatorname{div}(\nabla \Psi f)$ , and then alternatively run these processes on a small time interval. Furthermore, the transport process can be understood as a Wasserstein gradient flow of the potential energy. By adopting a suitable interpolation of the individual processes, they were able to show that the constructed splitting scheme converges to a weak solution of (1.3). In the literature, the technique of operator splitting is often used to construct

numerical methods for solving PDEs, see [27]. On the theoretical side, the idea of splitting had also been used to study the well-posedness of PDEs, see [9, 2] on kinetic equations and [3, 16] on fractional PDEs.

In the present work, we adopt the same splitting argument in [8] to construct a weak solution to the fractional kinetic equation (1.1). More specifically, we split the dynamics described in (1.1) by two phases:

- (1) Fractional diffusion phase. At every fixed position  $x \in \mathbf{R}^d$ , the probability density  $f(x, v, t)$ , as a function of velocity  $v$ , evolves according to the fractional heat equation

$$(1.4) \quad \partial_t f = -(-\Delta_v)^s f.$$

- (2) Kinetic transport phase. The density  $f(x, v, t)$  evolves according to the following equation

$$(1.5) \quad \partial_t f + v \cdot \nabla_x f = \operatorname{div}_v(\nabla \Psi(v) f).$$

We expect that successive alternative iterating the above two phases with vanishing period of time would give an approximation to the dynamics (1.1). The key difference between our splitting scheme above and the scheme in [8] is that the transport process here is not only driven by the potential energy but also the kinetic energy. In [8], the transport process is approximated by a discrete Wasserstein gradient flow based on the work [31]. However, due to the presence of the kinetic term, the kinetic transport equation is not a Wasserstein gradient flow; thus one can no longer use the Wasserstein distance. To overcome this obstacle, we employ instead the minimal acceleration cost function and the associated Kantorovich optimal transportation cost functional that has been used in [28, 19] for the kinetic Fokker-Planck equation and in [25] for the isentropic Euler system, see Section 3.

**1.1. Main result.** Throughout the paper, we make the following important assumption on the potential  $\Psi$ .

**Assumptions 1.1.**  $\Psi$  is non-negative and  $\Psi \in C^{1,1}(\mathbf{R}^d) \cap C^{2,1}(\mathbf{R}^d)$ .

We adopt the following notion of weak solution to KFPE (1.1).

**Definition 1.2.** Let  $f_0$  be a non-negative function such that  $f_0 \in \mathcal{P}_a^2(\mathbf{R}^{2d}) \cap L^p(\mathbf{R}^{2d})$  for some  $1 < p \leq \infty$  and  $\int_{\mathbf{R}^{2d}} f_0(x, v) \Psi(v) dv dx < \infty$ . We say that  $f(x, v, t)$  is a weak solution to (1.1) if it satisfies the following:

- (1)  $\int_{\mathbf{R}^{2d}} f(x, v, t) dx dv = \int_{\mathbf{R}^{2d}} f_0(x, v) dx dv = 1$  for any  $t \in (0, T)$ .
- (2)  $f(x, v, t) \geq 0$  for a.e.  $(x, v, t) \in \mathbf{R}^{2d} \times (0, T)$ .
- (3) For any test function  $\varphi \in C_c^\infty(\mathbf{R}^{2d} \times (-T, T))$ ,

$$\begin{aligned} & \int_0^T \int_{\mathbf{R}^{2d}} f(x, v, t) (\partial_t \varphi + v \cdot \nabla_x \varphi - \nabla_v \Psi \cdot \nabla_v \varphi - (-\Delta_v)^s \varphi) dt dx dv \\ & + \int_{\mathbf{R}^{2d}} f_0(x, v) \varphi(0, x, v) = 0. \end{aligned}$$

The main result of the paper is the following theorem.

**Theorem 1.3.** Suppose that Assumption 1.1 holds. Given a  $f_0 \in \mathcal{P}_a^2(\mathbf{R}^{2d}) \cap L^p(\mathbf{R}^{2d})$  for some  $1 < p \leq \infty$  and  $\int_{\mathbf{R}^{2d}} f_0(x, v) \Psi(v) dv dx < \infty$ , there exists a weak solution  $f(x, v, t)$  to (1.1) in the sense of Definition 1.2.

The proof of Theorem 1.3 is constructive, that is we will build a converging sequence to a solution of (1.1) from the splitting scheme discussed above that will be rigorously formulated in Section 4. The proof is based on a series of lemmas and is postponed to Section 5. As a by-product of the analysis, we also construct a discrete variational scheme and obtain its convergence for the kinetic transport

equation, see Theorem 3.3 in Section 3; thus extending the work [31] to include the kinetic feature. Furthermore, some possible extensions to more complex systems are discussed in Section 6. It is not clear to us how to obtain the uniqueness and regularity result. The bootstrap argument in [30] to prove smoothness of weak solutions (and hence also uniqueness) seems not working for the fractional Laplacian operator due to the lack of a product rule. It should be mentioned that in the recent paper [34], the author has proved the existence and uniqueness of a solution to the fractional Fokker-Planck equation (1.3) in some weighted Lebesgue spaces. It would be an interesting problem to generalize [34] to FKFPE. This is to be investigated in future work.

**1.2. Organization of the paper.** The rest of the paper is organized as follows. Section 2 summarizes some basic results about the fractional heat equation. Section 3 studies the kinetic transport equation and its variational formulation. The splitting scheme of the paper is formulated explicitly in Section 4 and some a priori estimates are established for the discrete sequences as well as their time-interpolation. The proof of the main result is presented in Section 5. Finally, in Section 6 we discuss several possible extensions of the analysis to more complex systems.

**1.3. Notation.** For  $k \in \mathbf{N}$ ,  $\alpha \in (0, 1]$ , let  $C^{k, \alpha}(\mathbf{R}^d)$  be the space of functions which have continuous derivative up to order  $k$  and whose  $k$ -th partial derivatives are Hölder continuous with exponent  $\alpha$ . Let  $\mathcal{P}^2(\mathbf{R}^d)$  be the collection of probability measures on  $\mathbf{R}^d$  with finite second moments. Let  $\mathcal{P}_a^2(\mathbf{R}^d)$  be the subset of probability measures in  $\mathcal{P}^2(\mathbf{R}^d)$  that are absolutely continuous with respect to the Lebesgue measure on  $\mathbf{R}^d$ . For  $\mu, \nu \in \mathcal{P}^2(\mathbf{R}^d)$ , the 2-Wasserstein distance  $W_2(\mu, \nu)$  is defined by

$$W_2(\mu, \nu) := \left( \inf \left\{ \int_{\mathbf{R}^{2d}} |x - y|^2 p(dx, dy) : p \in \mathcal{P}(\mu, \nu) \right\} \right)^{\frac{1}{2}}$$

where  $\mathcal{P}(\mu, \nu)$  is the set of probability measures on  $\mathbf{R}^{2d}$  with marginals  $\mu, \nu$ , i.e.  $p \in \mathcal{P}(\mu, \nu)$  if and only if

$$p(A \times \mathbf{R}^d) = \mu(A), \quad p(\mathbf{R}^d \times A) = \nu(A)$$

hold every Borel set  $A \in \mathbf{R}^d$ . In the case that  $\mu, \nu \in \mathcal{P}_a^2(\mathbf{R}^d)$  with densities  $f, g$ , we may write  $W_2(f, g)$  instead of  $W_2(\mu, \nu)$ .

We use the notation  $F^\# \mu$  to denote the push-forward of a probability measure  $\mu$  on  $\mathbf{R}^{2d}$  under map  $F$ , that is a probability measure on  $\mathbf{R}^{2d}$  satisfying for all smooth test function  $\varphi$ ,

$$\int_{\mathbf{R}^{2d}} \varphi(x, v) dF^\# \mu = \int_{\mathbf{R}^{2d}} \varphi(F(x, v)) d\mu.$$

## 2. THE FRACTIONAL HEAT EQUATION

This section collects some basic results on the fractional heat equation. We start by defining the fractional heat kernel

$$(2.1) \quad \Phi_s(v, t) := \mathcal{F}^{-1}(e^{-t|\cdot|^{2s}})(v).$$

Remember that the fractional Laplacian operator in (1.1) is only an operator in  $v$ -variable. With the fractional heat kernel, the solution to the fractional heat equation (1.4) with initial condition  $f_0(x, v)$  can be expressed as

$$(2.2) \quad f(x, v, t) = \Phi_s(\cdot, t) *_v f_0(x, v)$$

where  $*_v$  is the convolution operator in  $v$ -variable. The following elementary result is immediate from the definition of the kernel; see also [8].

**Lemma 2.1.**

- (1) For any  $t > 0$ ,  $\|\Phi_s(\cdot, t)\|_{L^1(\mathbf{R}^d)} = 1$ .
- (2) For any  $t > 0$  and  $p \in (1, \infty)$ ,  $\|\Phi_s(\cdot, t) * f_0\|_{L^p(\mathbf{R}^{2d})} \leq \|f_0\|_{L^p(\mathbf{R}^{2d})}$ .
- (3)  $\int_{\mathbf{R}} |v|^2 \Phi_s(v, t) dv = +\infty$  for all  $s \in (0, 1)$  and  $t > 0$ .

Lemma 2.1 (3) demonstrates a significant difference between the fractional heat kernel and standard Gaussian kernel, i.e. the former has infinite second moment. The loss of second moment bound may lead to infinite potential energy for example when the potential  $\Psi(v) = |v|^2$ . To overcome this issue, it is more convenient to make a renormalisation on the fractional heat kernel. To be more precise, for any  $h > 0$ , let us denote  $\Phi_s^h(v) := \Phi_s(v, h)$  and set  $\Phi_{s,R}^h(v) := \Phi_s^h(v) \mathbf{1}_{B_R}(v)$  where  $\mathbf{1}_{B_R}$  is an indicator function of a centred ball of radius  $R$ . Given a function  $f \in \mathcal{P}_a^2(\mathbf{R}^d)$ , we can define the renormalised convolution

$$(2.3) \quad \bar{f}_{h,R} := \frac{\Phi_{s,R}^h * f}{\|\Phi_{s,R}^h\|_{L^1(\mathbf{R}^d)}}$$

It is clear that the newly defined convolution satisfies  $\bar{f}_{h,R} \rightarrow \Phi_s^h * f$  pointwise. Moreover, we have the following lemma.

**Lemma 2.2.** *Let  $f$  be a function such that  $F \in C^{1,1}(\mathbf{R}^d) \cap C^2(\mathbf{R}^d)$ . Suppose that  $f \in \mathcal{P}_a^2(\mathbf{R}^{2d})$  and with  $\int_{\mathbf{R}^{2d}} f(x, v) F(v) dv dx < \infty$ . Then*

- (1)  $\bar{f}_{h,R} \in \mathcal{P}_a^2(\mathbf{R}^{2d})$ .
- (2)

$$(2.4) \quad \begin{aligned} \int_{\mathbf{R}^{2d}} \bar{f}_{h,R}(x, v) F(v) dx dv &\leq \int_{\mathbf{R}^{2d}} f(x, v) F(v) dx dv \\ &+ \frac{1}{2} \|D^2 F\|_{L^\infty(\mathbf{R}^d)} \frac{\int_{B_R} |w|^2 \Phi_s^h(w) dw}{\int_{B_R} \Phi_s^h(w) dw}. \end{aligned}$$

*Proof.* Notice that it suffices to prove part (2) since part (1) follows directly from part (2) by setting  $F(v) = |v|^2$ . The proof is similar to that of [8, Lemma 4.1], but for completeness we give the proof below. First from the definition of  $\bar{f}_{h,R}$ , one sees that

$$\int_{\mathbf{R}^{2d}} \bar{f}_{h,R}(x, v) F(v) dx dv = \frac{\int_{\mathbf{R}^{2d}} F(v) \int_{B_R} \Phi_s^h(w) f(x, v-w) dw dx dv}{\int_{B_R} \Phi_s^h(w) dw}.$$

Using change of variable  $z = v - w$  and Taylor's expansion, we can write the numerator as

$$\begin{aligned} &\int_{\mathbf{R}^{2d}} F(v) \int_{B_R} \Phi_s^h(w) f(x, v-w) dw dx dv \\ &= \int_{\mathbf{R}^{2d}} F(w+z) \int_{B_R} \Phi_s^h(w) f(x, z) dw dx dz \\ &= \int_{B_R} \Phi_s^h(w) \int_{\mathbf{R}^{2d}} F(w+z) f(x, z) dx dz dw \\ &= \int_{B_R} \Phi_s^h(w) \int_{\mathbf{R}^{2d}} \left[ F(z) + w \cdot \nabla F(z) + \frac{1}{2} w^T D^2 F(\xi_{w,z}) w \right] f(x, z) dx dz dw \\ &\leq \int_{B_R} \Phi_s^h(w) dw \int_{\mathbf{R}^{2d}} F(z) f(x, z) dx dz + \left| \int_{B_R} \Phi_s^h(w) \int_{\mathbf{R}^{2d}} w \cdot \nabla F(z) f(x, z) dx dz dw \right| \\ &\quad + \frac{1}{2} \|D^2 F\|_\infty \int_{B_R} |w|^2 \Phi_s^h(w) dw \int_{\mathbf{R}^{2d}} f(x, z) dx dz \\ &= \int_{B_R} \Phi_s^h(w) dw \int_{\mathbf{R}^{2d}} F(z) f(x, z) dx dz + \frac{1}{2} \|D^2 F\|_\infty \int_{B_R} |w|^2 \Phi_s^h(w) dw. \end{aligned}$$

Note that in the above  $\xi_{w,z}$  is an intermediate point between  $w$  and  $z$  and the term with the modulus vanishes since the kernel  $\Phi_s^h$  is symmetric with respect to the origin.  $\square$

The following lemma provides an upper bound for the ratio on the right side of (2.4).

**Lemma 2.3.** *For any  $s \in (0, 1]$ , there exists a constant  $C > 0$  such that*

$$(2.5) \quad \frac{\int_{B_R} |w|^2 \Phi_s^h(w) dw}{\int_{B_R} \Phi_s^h(w) dw} \leq C(h^{\frac{1}{s}} + hR^{2-2s})$$

holds for all  $R, h > 0$ .

*Proof.* This lemma follows directly from a two-sided point-wise estimate on  $\Phi_s^h(w)$  as shown in [8, Proposition 2.1]. See also equation (16) in [8].  $\square$

### 3. THE KINETIC TRANSPORT EQUATION AND ITS VARIATIONAL FORMULATION

**3.1. The minimum acceleration cost.** Consider the kinetic transport equation with initial value  $f_0$

$$(3.1) \quad \begin{aligned} \partial_t f(x, v, t) + v \cdot \nabla_x f &= \operatorname{div}_v(\nabla \Psi(v) f(x, v, t)), \\ f(x, v, 0) &= f_0(x, v). \end{aligned}$$

We are interested in the variational structure of (3.1) which is an interesting problem on its own right. In [31], Kinderlehrer and Tudorascu proved that the transport equation  $\partial_t f(x, v, t) = \operatorname{div}_v(\nabla \Psi(v) f)$ , which is the spatially homogeneous version of (3.1), is a Wasserstein gradient flow of the energy  $\int_{\mathbf{R}^d} \Psi f$ . Their proof is via constructing a discrete variational scheme as in [30]. However, due to the absence of the entropy term, which is super-linear, several non-trivial technicalities were introduced to obtain the compactness of the discrete approximations thus establishing the convergence of the scheme. For the kinetic transport equation (3.1), due to the presence of the kinetic term, it is not a Wasserstein gradient flow in the phase space thus the Wasserstein distance can no longer be used. Therefore to construct a discrete variational scheme for this equation, we need a different Kantorovich optimal transportation cost functional. To this end, we will employ the Kantorovich optimal transportation cost functional that is associated to the minimal acceleration cost. This cost functional has been used before in [28, 19] for the kinetic Fokker-Planck equation and in [25] for the isentropic Euler system. We follow the heuristics of defining the minimal acceleration cost as in [25]. Consider the motion of particle going from position  $x$  with velocity  $v$  to a new position  $x'$  with velocity  $v'$ , within a time interval of length  $h$ . Suppose that the particle follows a curve  $\xi : [0, h] \mapsto \mathbf{R}^d$  such that

$$(\xi, \dot{\xi})|_{t=0} = (x, v) \quad \text{and} \quad (\xi, \dot{\xi})|_{t=h} = (x', v')$$

and such that the average acceleration cost along the curve, that is  $\frac{1}{h} \int_0^h |\ddot{\xi}(t)|^2 dt$  is minimized. Then the curve is actually a cubic polynomial and the minimal average acceleration cost is given by  $C_h(x, v; x', v')/h^2$  where

$$(3.2) \quad C_h(x, v; x', v') := |v' - v|^2 + 12 \left| \frac{x' - x}{h} - \frac{v' + v}{2} \right|.$$

The Kantorovich functional  $\mathcal{W}_h(\mu, \nu)$  associated with the cost function  $C_h$ , is defined by, for any  $\mu, \nu \in \mathcal{P}^2(\mathbf{R}^{2d})$ ,

$$(3.3) \quad \mathcal{W}_h(\mu, \nu)^2 = \inf_{p \in \mathcal{P}(\mu, \nu)} \int_{\mathbf{R}^{4d}} C_h(x, v; x', v') p(dx dv dx' dv'),$$



where  $\mathcal{P}(\mu, \nu)$  is the set of all couplings between  $\mu$  and  $\nu$ . It is important to notice that  $\mathcal{W}_h$  is not a distance. In fact,  $\mathcal{W}_h$  is not symmetric in the arguments  $\mu, \nu$ , due to the asymmetry of the cost function  $C_h$ . In addition,  $\mathcal{W}_h(\mu, \nu)$  does not vanish when  $\mu = \nu$ . Instead, we have that

$$\mathcal{W}_h(\mu, \nu) = 0 \iff \nu = (F_h)_\# \mu,$$

where  $F_h$  is the free transport map defined by

$$(3.4) \quad \begin{aligned} F_h : \mathbf{R}^d \times \mathbf{R}^d &\rightarrow \mathbf{R}^d \times \mathbf{R}^d \\ (x, v) &\mapsto F_h(x, v) = (x + hv, v). \end{aligned}$$

It is also useful to define the map

$$(3.5) \quad G_h : \mathbf{R}^d \times \mathbf{R}^d \rightarrow \mathbf{R}^d \times \mathbf{R}^d \\ (x, v) \mapsto G_h(x, v) = \left( \sqrt{3} \left( \frac{2x}{h} - v \right), v \right).$$

The composition  $G_h \circ F_h$  is then given by

$$(G_h \circ F_h)(x, v) = \left( \sqrt{3} \left( \frac{2x}{h} + v \right), v \right).$$

The motivation of introducing the map  $G_h$  comes from the following identity which links the Kantorovich functional  $\mathcal{W}_h$  for two Dirac mass and the standard 2-Wasserstein distance of their pushforwards:

$$\begin{aligned} \mathcal{W}_h(\delta_{(x,v)}, \delta_{(x',v')})^2 &= C_h(x, v; x', v') \\ &= \left| G_h \circ F_h(x, v) - G_h(x', v') \right|^2 \\ &= W_2((G_h \circ F_h)_\# \delta_{(x,v)}, G_h^\# \delta_{(x',v')})^2. \end{aligned}$$

The next lemma formalized the above link in general cases, whose proof is similar to that of [25, Proposition 4.4].

**Lemma 3.1.** [25, Proposition 4.4] *Let  $F_h$  and  $G_h$  be given by (3.4) and (3.5) respectively. The Kantorovich functional  $\mathcal{W}_h$  can be expressed in terms of the 2-Wasserstein distance  $W_2$  as*

$$(3.6) \quad \mathcal{W}_h(\mu, \nu) = W_2((G_h \circ F_h)_\# \mu, G_h^\# \nu) \quad \text{for all } \mu, \nu \in \mathcal{P}^2(\mathbf{R}^{2d}).$$

As a consequence, the infimum in (3.3) is attained and thus  $\mathcal{W}_h(\mu, \nu)$  is a minimum.

**3.2. Variational formulation.** With  $\mathcal{W}_h$  being defined, we want to interpret (3.1) as a generalized gradient flow of the potential energy  $\int_{\mathbf{R}^{2d}} \Psi(v) f(x, v) dx dv$  with respect to  $\mathcal{W}_h$ . For doing so, we consider the variational problem

$$(3.7) \quad \inf_{f \in \mathcal{P}_a^2} \mathcal{A}(f) := \frac{1}{2h} \mathcal{W}_h(f_0, f)^2 + \int_{\mathbf{R}^{2d}} \Psi(v) f(x, v) dx dv.$$

Here  $f_0 \in \mathcal{P}_a^2(\mathbf{R}^{2d})$  is an initial probability density with  $\int_{\mathbf{R}^{2d}} \Psi(v) f_0(x, v) dx dv < \infty$  and  $h > 0$  is the time step. The next lemma establishes some properties about the minimizer to (3.7).

**Lemma 3.2.**

- (1) *For  $h$  being sufficiently small, the variational problem (3.7) has a unique minimizer  $f \in \mathcal{P}_a^2(\mathbf{R}^{2d})$ .*
- (2) *Let  $h > 0$  be small enough such that  $\det(I + hD^2(\Psi(v))) \leq 1 + \alpha h$  for some fixed  $\alpha > \|D^2\Psi\|_{L^\infty(\mathbf{R}^d)}$ . If  $f_0 \in L^p(\mathbf{R}^{2d})$  for  $1 < p < \infty$ , then*

$$(3.8) \quad \|f\|_{L^p(\mathbf{R}^{2d})}^p \leq (1 - \alpha h)^{p-1} \|f_0\|_{L^p(\mathbf{R}^{2d})}^p.$$



(3)  $f$  satisfies the following Euler-Lagrange equation: for any  $\varphi \in C_c^\infty(\mathbf{R}^{2d})$ ,

$$(3.9) \quad \begin{aligned} & \frac{1}{h} \int_{\mathbf{R}^{4d}} [(x' - x) \cdot \nabla_{x'} \varphi(x', v') + (v' - v) \cdot \nabla_{v'} \varphi(x', v')] P^*(dx dv dx' dv') \\ & - \int_{\mathbf{R}^{2d}} v' \cdot \nabla_{x'} \varphi(x', v') f(x', v') dx' dv' \\ & + \int_{\mathbf{R}^{2d}} \nabla_{v'} \Psi(v') \cdot \nabla_{v'} \varphi(x', v') f(x', v') dx' dv' = \mathcal{R}, \end{aligned}$$

where  $P^*$  is the optimal coupling in  $\mathcal{W}_h(f_0, f)$  and

$$(3.10) \quad \begin{aligned} \mathcal{R} &= -\frac{h}{2} \int_{\mathbf{R}^{4d}} \nabla_{v'} \Psi(v') \cdot \nabla_{x'} \varphi(x', v') P^*(dx dv dx' dv') \\ &= -\frac{h}{2} \int_{\mathbf{R}^{2d}} \nabla_{v'} \Psi(v') \cdot \nabla_{x'} \varphi(x', v') f(x', v') dx' dv'. \end{aligned}$$

*Proof.*

(1) Thanks to Lemma 3.1, we can rewrite the functional  $\mathcal{A}$  as

$$\begin{aligned} \mathcal{A}(f) &= \frac{1}{2h} W_2((G_h \circ F_h)^\# f_0, (G_h)^\# f)^2 + \int_{\mathbf{R}^{2d}} \Psi(v) (G_h)^\# f(dx dv) \\ &= \frac{1}{2h} W_2(\tilde{f}_0, \tilde{f})^2 + \int_{\mathbf{R}^{2d}} \Psi(v) \tilde{f}(dx dv) =: \tilde{A}(\tilde{f}), \end{aligned}$$

where  $\tilde{f}_0 = (G_h \circ F_h)^\# f_0$  and  $\tilde{f} = (G_h)^\# f$ . According to [31, Proposition 1] (see also [8, Proposition 3.1]), the functional  $\tilde{A}$  has a unique minimizer, denoted by  $\tilde{f}$ . Therefore, the problem (3.7) has a unique minimizer  $f = (G_h^{-1})^\# \tilde{f}$ .

(2) This follows directly from [31, Proposition 1] and the fact that if  $\tilde{f} = (G_h)^\# f$  then

$$\|\tilde{f}\|_{L^p(\mathbf{R}^{2d})}^p = \left( \frac{2h}{\sqrt{3}} \right)^{d(p-1)} \|f\|_{L^p(\mathbf{R}^{2d})}^p.$$

(3) The derivation of the Euler-Lagrange equation for the minimizer  $f$  of the variational problem (3.7) follows the now well-established procedure (see e.g. [30, 28]). For the reader's convenience, we sketch the main steps here. First, we consider the perturbation of  $f$  defined by push-forwarding  $f$  under the flows  $\phi, \psi: [0, \infty) \times \mathbf{R}^{2d} \rightarrow \mathbf{R}^d$ :

$$\begin{aligned} \frac{\partial \psi_s}{\partial s} &= \zeta(\psi_s, \phi_s), \quad \frac{\partial \phi_s}{\partial s} = \eta(\psi_s, \phi_s), \\ \psi_0(x, v) &= x, \quad \phi_0(x, v) = v, \end{aligned}$$

where  $\zeta, \eta \in C_0^\infty(\mathbf{R}^{2d}, \mathbf{R}^d)$  will be chosen later. Let us denote  $\gamma_s$  to be the push forward of  $f$  under the flow  $(\psi_s, \phi_s)$ . Since  $(\psi_0, \phi_0) = \text{Id}$ , it follows that  $\gamma_0 = f$ , and an explicit calculation gives

$$(3.11) \quad \partial_s \gamma_s|_{s=0} = -\text{div}_x(f\zeta) - \text{div}_v(f\eta)$$

in the sense of distributions. Second, thanks to the optimality of  $f$ , we have that  $\mathcal{A}(\gamma_s) \geq \mathcal{A}(f)$  for all  $\gamma_s$  defined via the flow above. Then a standard variational argument as in [30, 28] leads to the following stationary equation on  $f$ :

$$(3.12) \quad \begin{aligned} & \frac{1}{2h} \int_{\mathbf{R}^{4d}} [\nabla_{x'} C_h(x, v; x', v') \cdot \zeta(x', v') + \nabla_{v'} C_h(x, v; x', v') \cdot \eta(x', v')] P^*(dx dv dx' dv') \\ & + \int_{\mathbf{R}^{2d}} f(x, v) \nabla_v \Psi(v) \cdot \eta(x, v) dx dv = 0, \end{aligned}$$

where  $\mathcal{P}^*$  is the optimal coupling in the definition of  $\mathcal{W}_h(f_0, f)$ . Third, we choose  $\zeta$  and  $\eta$  with a given  $\varphi \in C_0^\infty(\mathbf{R}^{2d}, \mathbf{R})$  as follows

$$(3.13) \quad \begin{aligned} \zeta(x', v') &= -\frac{h^2}{6} \nabla_{x'} \varphi(x', v') + \frac{1}{2} h \nabla_{v'} \varphi(x', v'), \\ \eta(x', v') &= -\frac{1}{2} h \nabla_{x'} \varphi(x', v') + \nabla_{v'} \varphi(x', v'). \end{aligned}$$

Now from the definition of the cost functional  $C_h(x, v; x', v')$  in (3.2), we have that

$$\begin{aligned} \nabla_{x'} C_h &= \frac{24}{h} \left( \frac{x' - x}{h} - \frac{v' + v}{2} \right), \\ \nabla_{v'} C_h &= 2(v' - v) - 12 \left( \frac{x' - x}{h} - \frac{v' + v}{2} \right). \end{aligned}$$

Therefore, together with (3.13), we calculate

$$\begin{aligned} &\nabla_{x'} C_h \cdot \zeta + \nabla_{v'} C_h \cdot \eta \\ &= \frac{24}{h} \left( \frac{x' - x}{h} - \frac{v' + v}{2} \right) \cdot \left( -\frac{h^2}{6} \nabla_{x'} \varphi(x', v') + \frac{1}{2} h \nabla_{v'} \varphi(x', v') \right) \\ &\quad + \left( 2(v' - v) - 12 \left( \frac{x' - x}{h} - \frac{v' + v}{2} \right) \right) \cdot \left( -\frac{1}{2} h \nabla_{x'} \varphi(x', v') + \nabla_{v'} \varphi(x', v') \right) \\ &= 2((x' - x) - hv') \cdot \nabla_{x'} \varphi + 2(v' - v) \cdot \nabla_{v'} \varphi. \end{aligned}$$

The Euler-Lagrange equation (3.9) for the minimizer  $f$  follows directly by substituting the equation above back into (3.12).  $\square$

We now can build up a discrete variational scheme for the kinetic transport equation as follows. Given  $f_0 \in \mathcal{P}_a^2(\mathbf{R}^{2d})$  with  $\int_{\mathbf{R}^{2d}} \Psi(v) f_0(x, v) dx dv < \infty$  and  $h > 0$  is the time step. For every integer  $k \geq 1$ , we define  $f_k$  as the minimizer of the minimization problem

$$(3.14) \quad \inf_{f \in \mathcal{P}_a^2} \left\{ \frac{1}{2h} \mathcal{W}_h(f_{k-1}, f)^2 + \int_{\mathbf{R}^{2d}} \Psi(v) f(x, v) dx dv \right\}.$$

The following theorem extends the work [31] to the kinetic transport equation.

**Theorem 3.3.** *Suppose that Assumption 1.1 holds. Given a  $f_0 \in \mathcal{P}_a^2(\mathbf{R}^{2d}) \cap L^p(\mathbf{R}^{2d})$  for some  $1 < p \leq \infty$  and  $\int_{\mathbf{R}^{2d}} f_0(x, v) \Psi(v) dv dx < \infty$ , there exists a weak solution  $f(x, v, t)$  to equation (3.1) in the sense of Definition 1.2 but with the fractional Laplacian term removed.*

*Proof.* The proof of this theorem follows the same lines as that of the Theorem 1.3, that is to show that the discrete variational scheme (3.14) above converges to a weak solution of the kinetic transport equation. Since the proof of Theorem 1.3 will be carried out in details in Section 5, we omit this proof here.  $\square$

#### 4. A SPLITTING SCHEME FOR FKFPE

**4.1. Definition of splitting scheme.** As we mentioned in the introduction section, we aim to construct an operator splitting scheme for equation (1.1) by continuously alternating processes (1.4) and (1.5), where the later is approximated by the generalized gradient flow of the potential energy, or equivalently, the density after a short time step  $h$  is approximately given by the solution to the variational problem (3.7). However, there is an issue associated with iterating (1.4) and (3.7). That is, the solution of the fractional heat equation may not have a finite second moment (see Lemma 2.1 (3)). Hence it can not be used as the initial condition in the variational problem (3.7) since the potential energy might be infinite. To around this issue, we define an approximate fractional diffusion process by using

the renormalised convolution (2.3) based on the truncated fractional heat kernel. To be more precise, given a fixed  $N \in \mathbb{N}$ , let us consider a uniform partition  $0 = t_0 < t_1 < \dots < t_N = T$  of the time interval  $[0, T]$  with  $t_k = kh$  and  $h = T/N$ . With an initial condition  $f_h^0 = f_0$ , for  $n = 1, \dots, N$  we iteratively compute the following:

- Given a truncation parameter  $R > 0$ , compute the renormalised convolution

$$(4.1) \quad \bar{f}_{h,R}^n := \frac{\Phi_{s,R}^h *_v f_{h,R}^{n-1}}{\|\Phi_{s,R}^h\|_{L^1(\mathbf{R}^d)}}.$$

- Solve for the minimizer  $f_{h,R}^n$  of the problem

$$(4.2) \quad f_{h,R}^n := \operatorname{argmin}_{f \in \mathcal{P}_a^2(\mathbf{R}^d)} \left\{ \frac{1}{2h} \mathcal{W}_h(\bar{f}_{h,R}^n, f)^2 + \int_{\mathbf{R}^{2d}} \Psi(v) f(x, v) dx dv \right\}.$$

Note that thanks to Lemma 3.2 (1) the minimizer  $f_{h,R}^n$  in (4.2) is well-defined and unique. Moreover, it follows from Lemma 3.2 (3) that  $f_{h,R}^n$  satisfies the following equation

$$(4.3) \quad \begin{aligned} & \int_{\mathbf{R}^{4d}} [(x' - x) \cdot \nabla_{x'} \varphi(x', v') + (v' - v) \cdot \nabla_{v'} \varphi(x', v')] P_{h,R}^n(dx dv dx' dv') \\ &= h \int_{\mathbf{R}^{2d}} (v' \cdot \nabla_{x'} \varphi(x', v') - \nabla_{v'} \Psi(v') \cdot \nabla_{v'} \varphi(x', v')) f_{h,R}^n(x', v') dx' dv' + \mathcal{R}_{h,R}^n, \end{aligned}$$

where  $P_{h,R}^n$  is the optimal coupling in  $\mathcal{W}_h(\bar{f}_{h,R}^n, f_{h,R}^n)$  and

$$(4.4) \quad \mathcal{R}_{h,R}^n = \frac{h^2}{2} \int_{\mathbf{R}^{2d}} \nabla_v \Psi(v) \cdot \nabla_x \varphi(x, v) f_{h,R}^n(dx dv).$$

With the scheme being defined above, we obtain a discrete approximating sequence  $\{f_{h,R}^n\}_{0 \leq n \leq N}$ . Below we define a time-interpolation based on  $\{f_{h,R}^n\}$  and our ultimate goal is to prove that this sequence converges to a weak solution of (1.1).

**Time-interpolation:** We define  $f_{h,R}$  by setting

$$(4.5) \quad f_{h,R}(t) := \Phi_s(t - t_n) *_v f_{h,R}^n \text{ for } t \in [t_n, t_{n+1}).$$

It is clear that by definition  $f_{h,R}$  solves the fractional heat equation on every interval  $[t_n, t_{n+1})$  with initial condition  $f_{h,R}^n$ . Notice also that  $f_{h,R}$  is only right-continuous in general. For convenience, we also define

$$(4.6) \quad \tilde{f}_{h,R}^{n+1} = \lim_{t \uparrow t_{n+1}} f_{h,R}(t).$$

**4.2. A priori estimates.** In this section, we prove some useful a priori estimates for the discrete-time sequence  $\{f_{h,R}^n\}$  as well as for the time-interpolation sequence  $\{f_{h,R}(t)\}$ . We start by proving an upper bound for the sum of the Kantorovich functionals  $\mathcal{W}_h(\bar{f}_{h,R}^n, f_{h,R}^n)$ .

**Lemma 4.1.** *Let  $\{\bar{f}_{h,R}^n\}$  and  $\{f_{h,R}^n\}$  be the sequences constructed from the splitting scheme (4.1)-(4.2). Then there exists a constant  $C > 0$ , independent of  $h$  and  $R$ , such that*

$$(4.7) \quad \sum_{n=1}^N \mathcal{W}_h(\bar{f}_{h,R}^n, f_{h,R}^n)^2 \leq C \left( h \int_{\mathbf{R}^{2d}} \Psi(v) f_0(x, v) dx dv + T \|D^2 \Psi\|_{\infty} (h^{1/2} + hR^{2-2s}) \right).$$

*Proof.* Since  $f_{h,R}^n$  minimizes the functional  $f \mapsto \frac{1}{2h} \mathcal{W}_h(\bar{f}_{h,R}^n, f) + \int_{\mathbf{R}^{2d}} \Psi(v) f(x, v) dx dv$ , for all  $f \in \mathcal{P}_2(\mathbf{R}^{2d})$ , we have

$$\frac{1}{2h} \mathcal{W}_h(\bar{f}_{h,R}^n, f_{h,R}^n)^2 + \int_{\mathbf{R}^{2d}} \Psi(v) f_{h,R}^n dx dv \leq \frac{1}{2h} \mathcal{W}_h(\bar{f}_{h,R}^n, f)^2 + \int_{\mathbf{R}^{2d}} \Psi(v) f(x, v) dx dv.$$

In particular, if we set  $f = f^* := F_h^\# \bar{f}_{h,R}^n$  where  $F_h$  is the free transport map defined in (3.4), then since  $\mathcal{W}_h(\bar{f}_{h,R}^n, f^*) = 0$  we obtain

$$\begin{aligned} \mathcal{W}_h(\bar{f}_{h,R}^n, f_{h,R}^n)^2 &\leq 2h \left( \int_{\mathbf{R}^{2d}} \Psi(v) f^*(x, v) dx dv - \int_{\mathbf{R}^{2d}} \Psi(v) f_{h,R}^n dx dv \right) \\ (4.8) \quad &= 2h \left( \int_{\mathbf{R}^{2d}} \Psi(v) \bar{f}_{h,R}^n(x, v) dx dv - \int_{\mathbf{R}^{2d}} \Psi(v) f_{h,R}^n dx dv \right). \end{aligned}$$

We have also used the fact that the free transport map  $F_h$  has unit Jacobian in the last equality. According to Lemma 2.2 (2), we have

$$\begin{aligned} \int_{\mathbf{R}^{2d}} \Psi(v) \bar{f}_{h,R}^n(x, v) dx dv &\leq \int_{\mathbf{R}^{2d}} \Psi(v) f_{h,R}^{n-1}(x, v) dx dv \\ (4.9) \quad &+ \frac{1}{2} \|D^2 \Psi\|_\infty \frac{\int_{B_R} |w|^2 \Phi_s^h(w) dw}{\int_{B_R} \Phi_s^h(w) dw}. \end{aligned}$$

Substituting (4.9) into (4.8), we obtain

$$\begin{aligned} \mathcal{W}_h(\bar{f}_{h,R}^n, f_{h,R}^n)^2 &\leq 2h \left( \int_{\mathbf{R}^{2d}} \Psi(v) f_{h,R}^{n-1}(x, v) dx dv - \int_{\mathbf{R}^{2d}} \Psi(v) f_{h,R}^n dx dv \right) \\ &+ h \|D^2 \Psi\|_\infty \frac{\int_{B_R} |w|^2 \Phi_s^h(w) dw}{\int_{B_R} \Phi_s^h(w) dw}, \end{aligned}$$

from which, by summing over  $n$  from 1 to  $N$  we obtain

$$(4.10) \quad \sum_{n=1}^N \mathcal{W}_h(\bar{f}_{h,R}^n, f_{h,R}^n)^2 \leq 2h \int_{\mathbf{R}^{2d}} \Psi(v) f_0(x, v) dx dv + T \|D^2 \Psi\|_\infty \frac{\int_{B_R} |w|^2 \Phi_s^h(w) dw}{\int_{B_R} \Phi_s^h(w) dw}.$$

Then the desired estimate follows directly from (4.10) and Lemma 2.3.  $\square$

We also need some second moment bounds on  $f$  with respect to variable  $v$ . Given a density function  $f$ , let us set  $M_{2,v}(f) := \int |v|^2 f(x, v) dx dv$ .

**Lemma 4.2.** *There exist positive constants  $C, h_0$  such that when  $0 < h < h_0$ , it holds for any index  $i > 0$  that*

$$(4.11) \quad M_{2,v}(f_{h,R}^i) \leq M_{2,v}(f_{h,R}^{i-1}) + 4\mathcal{W}_h(\bar{f}_{h,R}^i, f_{h,R}^i)^2 + C(h^{1/s} + hR^{2-2s}).$$

It follows that

$$\begin{aligned} (4.12) \quad &\max \left\{ M_{2,v}(f_{h,R}^n), M_{2,v}(\bar{f}_{h,R}^n) \right\} \leq M_{2,v}(f^0) \\ &+ 4 \sum_{i=1}^n \mathcal{W}_h(\bar{f}_{h,R}^i, f_{h,R}^i)^2 + C(n+1)(h^{1/s} + hR^{2-2s}). \end{aligned}$$

In addition, we have

$$\begin{aligned} (4.13) \quad &\int \left( |x - x'|^2 + |v - v'|^2 \right) P_{h,R}^i(dx dv dx' dv') \leq C\mathcal{W}_h(\bar{f}_{h,R}^i, f_{h,R}^i)^2 \\ &+ Ch^2 \left( M_{2,v}(\bar{f}_{h,R}^i) + M_{2,v}(f_{h,R}^i) \right), \end{aligned}$$

where  $P_{h,R}^i$  denotes the optimal coupling in the definition of  $\mathcal{W}_h(\bar{f}_{h,R}^i, f_{h,R}^i)$ .

*Proof.* First from the definition of the cost function  $C_h$  in (3.2) we have the following inequalities:

$$\begin{aligned}
 (4.14) \quad & |v' - v|^2 \leq C_h(x, v; x', v'); \\
 & |x' - x|^2 = h^2 \left| \frac{x' - x}{h} - \frac{v' + v}{2} + \frac{v' + v}{2} \right|^2 \\
 & \leq h^2 \left[ 2 \left| \frac{x' - x}{h} - \frac{v' + v}{2} \right|^2 + \frac{|v' + v|^2}{2} \right] \\
 (4.15) \quad & \leq h^2 \left( \frac{1}{6} C_h(x, v; x', v') + \frac{|v'|^2 + |v|^2}{2} \right).
 \end{aligned}$$

Then there exist constants  $C, h_0 > 0$  such that when  $h < h_0$ ,

$$(4.16) \quad |x' - x|^2 + |v' - v|^2 \leq C C_h(x, v; x', v') + h^2(|v'|^2 + |v|^2).$$

Now for any fixed  $i > 0$ , we have

$$\begin{aligned}
 \int_{\mathbf{R}^{2d}} |v|^2 f_{h,R}^i &= \int_{\mathbf{R}^{4d}} |v'|^2 P_{h,R}^i(dx dv dx' dv') \\
 &\leq \int_{\mathbf{R}^{4d}} |v' - v|^2 P_{h,R}^i(dx dv dx' dv') + \int_{\mathbf{R}^{4d}} |v|^2 P_{h,R}^i(dx dv dx' dv') \\
 &\stackrel{(4.14)}{\leq} 4 \mathcal{W}_h(\bar{f}_{h,R}^i, f_{h,R}^i)^2 + \int_{\mathbf{R}^{2d}} |v|^2 \bar{f}_{h,R}^i dx dv \\
 &\stackrel{(2.4)}{\leq} 4 \mathcal{W}_h(\bar{f}_{h,R}^i, f_{h,R}^i)^2 + \int_{\mathbf{R}^{2d}} |v|^2 f_{h,R}^{i-1} dx dv + C(h^{1/s} + hR^{2-2s}).
 \end{aligned}$$

This proves (4.11). The estimate (4.12) follows by summing the estimate (4.11) over the index  $i$  from 1 to  $n$  and inequality (2.4) with  $F(v) = |v|^2$ . Finally, the estimate (4.13) follows directly from inequality (4.16) and the definition of  $\mathcal{W}_h(\bar{f}_{h,R}^i, f_{h,R}^i)$ .  $\square$

In the next lemma, we prove a uniform  $L^p$ -bound for the time-interpolation sequence  $\{f_{h,R}\}$ .

**Lemma 4.3.** *Let  $h > 0$  be small enough such that  $\det(I + hD^2(\Psi(v))) \leq 1 + \alpha h$  for some fixed  $\alpha > \|D^2\Psi\|_{L^\infty(\mathbf{R}^d)}$ . If  $f_0 \in L^p(\mathbf{R}^{2d})$  for  $1 < p < \infty$ , then*

$$(4.17) \quad \|f_{h,R}(t)\|_{L^p(\mathbf{R}^{2d})}^p \leq e^{\alpha T(1-p)} \|f_0\|_{L^p(\mathbf{R}^{2d})}^p.$$

*Proof.* First, according to Lemma 3.2 (2), we have that

$$\|f_{h,R}^n\|_{L^p(\mathbf{R}^{2d})}^p \leq (1 - \alpha h)^{p-1} \|\bar{f}_{h,R}^n\|_{L^p(\mathbf{R}^{2d})}^p.$$

In addition, by the definition of  $\bar{f}_{h,R}^n$  (see (4.1)) and Young's inequality for convolution,

$$\|\bar{f}_{h,R}^n\|_{L^p(\mathbf{R}^{2d})}^p \leq \|f_{h,R}^{n-1}\|_{L^p(\mathbf{R}^{2d})}^p.$$

This implies that for any  $n > 0$ ,

$$\|f_{h,R}^n\|_{L^p(\mathbf{R}^{2d})}^p \leq (1 - \alpha h)^{n(p-1)} \|f_0\|_{L^p(\mathbf{R}^{2d})}^p.$$

Then by the definition of the time-interpolation  $f_{h,R}$  in (4.5), we have for any  $t \in (t_n, t_{n+1})$  that

$$\begin{aligned}
 \|f_{h,R}(t)\|_{L^p(\mathbf{R}^{2d})}^p &= \|\Phi_s(t - t_n) *_v f_{h,R}^n\|_{L^p(\mathbf{R}^{2d})}^p \\
 &\leq \|f_{h,R}^n\|_{L^p(\mathbf{R}^{2d})}^p \\
 &\leq (1 - \alpha h)^{n(p-1)} \|f_0\|_{L^p(\mathbf{R}^{2d})}^p \\
 &\leq e^{\alpha T(1-p)} \|f_0\|_{L^p(\mathbf{R}^{2d})}^p.
 \end{aligned}$$

□

## 5. PROOF OF THEOREM 1.3

In this section we prove the main result (Theorem 1.3) of the present paper. The major technical difficulty is to show that the time-interpolation  $f_{h,R}$  defined in (4.5) satisfies an approximate equation to the FKFPE with an error term that vanishes, with a suitable choice of the truncated parameter  $R$ , when  $h$  tends to zero. The key observation is that, as the minimizer of the variational problem (4.2),  $f_{h,R}^n$  satisfies an approximate equation to the kinetic transport equation (cf. Lemma 3.2). Combining this with the fact that, by definition,  $f_{h,R}$  solves the fractional heat equation on every interval  $[t_n, t_{n+1})$  with initial condition  $f_{h,R}^n$  will lead to the desired result.

**5.1. Approximate equation.** We first show in the next lemma that the time-interpolation  $f_{h,R}$  satisfies an approximate equation.

**Lemma 5.1.** *Let  $\varphi \in C_c^\infty([0, T] \times \mathbf{R}^d \times \mathbf{R}^d)$  with time support in  $[-T, T]$ . Then*

$$(5.1) \quad \int_0^T \int_{\mathbf{R}^{2d}} f_{h,R} [\partial_t \varphi + v \cdot \nabla_x \varphi - \nabla_v \Psi \cdot \nabla_v \varphi - (-\Delta_v)^s \varphi] dx dv dt + \int_{\mathbf{R}^{2d}} f_0(x, v) \varphi(0, x, v) dx dv = \mathcal{R}(h, R),$$

where  $\mathcal{R}(h, R) = \sum_{j=1}^4 \mathcal{R}_j(h, R) + \tilde{\mathcal{R}}(h, R)$  and

(5.2)

$$\mathcal{R}_1(h, R) = \sum_{n=1}^N \int_{\mathbf{R}^{2d}} \varphi(t_n) (\tilde{f}_{h,R}^n - \bar{f}_{h,R}^n) dx dv,$$

$$(5.3) \quad \mathcal{R}_2(h, R) = \sum_{n=1}^{N-1} \int_{t_n}^{t_{n+1}} \int_{\mathbf{R}^{2d}} \left( (v \cdot \nabla_x \varphi(t, x, v) - \nabla_v F(v) \cdot \nabla_v \varphi(t, x, v)) f_{h,R}(t, x, v) - (v \cdot \nabla_x \varphi(t_n, x, v) - \nabla_v F(v) \cdot \nabla_v \varphi(t_n, x, v)) f_{h,R}^n(x, v) \right) dx dv dt,$$

(5.4)

$$\mathcal{R}_3(h, R) = \int_0^h \int_{\mathbf{R}^{2d}} \Phi_s(t) * f_0 (v \cdot \nabla_x \varphi(t, x, v) - \nabla_v F(v) \cdot \nabla_v \varphi(t, x, v)) dx dv dt,$$

(5.5)

$$\mathcal{R}_4(h, R) = \frac{h^2}{2} \sum_{n=1}^N \int_{\mathbf{R}^2} \nabla_v \Psi(v) \cdot \nabla_x \varphi(x, v) f_{h,R}^n(dx dv).$$

Moreover,

$$\tilde{\mathcal{R}}(h, R) \leq \frac{1}{2} \sum_{n=1}^N \|\nabla^2 \varphi(t_n)\|_\infty \int_{\mathbf{R}^4} (|x - x'|^2 + |v - v'|^2) P_{h,R}^n(dx dv dx' dv').$$

Here  $P_{h,R}^n$  is the optimal coupling in the definition of  $\mathcal{W}_h(\bar{f}_{h,R}^n, f_{h,R}^n)$ .

*Proof.* From the definition of  $f_{h,R}$  (see (4.5)) and integration by parts, we obtain that

$$\begin{aligned}
 (5.6) \quad & \int_{t_n}^{t_{n+1}} \int_{\mathbf{R}^{2d}} f_{h,R}(t) \partial_t \varphi(t) dt dx dv \\
 &= \int_{\mathbf{R}^{2d}} (\varphi(t_{n+1}) \tilde{f}_{h,R}^{n+1} - \varphi(t_n) f_{h,R}^n) dx dv - \int_{t_n}^{t_{n+1}} \int_{\mathbf{R}^{2d}} \varphi(t) \partial_t f_{h,R}(t) dt dx dv \\
 &= \int_{\mathbf{R}^{2d}} (\varphi(t_{n+1}) \tilde{f}_{h,R}^{n+1} - \varphi(t_n) f_{h,R}^n) dx dv + \int_{t_n}^{t_{n+1}} \int_{\mathbf{R}^{2d}} \varphi(t) (-\Delta_v)^s f_{h,R}(t) dt dx dv \\
 &= \int_{\mathbf{R}^{2d}} (\varphi(t_{n+1}) \tilde{f}_{h,R}^{n+1} - \varphi(t_n) f_{h,R}^n) dx dv + \int_{t_n}^{t_{n+1}} \int_{\mathbf{R}^{2d}} f_{h,R}(t) (-\Delta_v)^s \varphi(t) dt dx dv,
 \end{aligned}$$

where the second equality holds because  $f_{h,R}$  solves the fractional heat equation.

By adding and subtracting a few terms, we can write the first term on the right hand side of (5.6) as

$$\begin{aligned}
 (5.7) \quad & \int_{\mathbf{R}^{2d}} (\varphi(t_{n+1}) \tilde{f}_{h,R}^{n+1} - \varphi(t_n) f_{h,R}^n) dx dv \\
 &= \int_{\mathbf{R}^{2d}} (\varphi(t_{n+1}) f_{h,R}^{n+1} - \varphi(t_n) f_{h,R}^n) dx dv + \int_{\mathbf{R}^{2d}} \varphi(t_{n+1}) (\tilde{f}_{h,R}^{n+1} - f_{h,R}^{n+1}) dx dv \\
 &= \int_{\mathbf{R}^{2d}} (\varphi(t_{n+1}) f_{h,R}^{n+1} - \varphi(t_n) f_{h,R}^n) dx dv + \int_{\mathbf{R}^{2d}} \varphi(t_{n+1}) (\tilde{f}_{h,R}^{n+1} - \bar{f}_{h,R}^{n+1}) dx dv \\
 &\quad + \int_{\mathbf{R}^{2d}} \varphi(t_{n+1}) (\bar{f}_{h,R}^{n+1} - f_{h,R}^{n+1}) dx dv.
 \end{aligned}$$

Now substituting (5.7) back into (5.6) and then summing over index  $n$  from 0 to  $N-1$  yields

$$\begin{aligned}
 (5.8) \quad & \int_0^T \int_{\mathbf{R}^{2d}} f_{h,R}(t) \partial_t \varphi(t) dt dx dv \\
 &= \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \int_{\mathbf{R}^{2d}} f_{h,R}(t) \partial_t \varphi(t) dt dx dv \\
 &= \sum_{n=0}^{N-1} \left[ \int_{t_n}^{t_{n+1}} \int_{\mathbf{R}^{2d}} f_{h,R}(t) (-\Delta_v)^s \varphi(t) dt dx dv \right. \\
 &\quad + \int_{\mathbf{R}^{2d}} (\varphi(t_{n+1}) f_{h,R}^{n+1} - \varphi(t_n) f_{h,R}^n) dx dv + \int_{\mathbf{R}^{2d}} \varphi(t_{n+1}) (\tilde{f}_{h,R}^{n+1} - \bar{f}_{h,R}^{n+1}) dx dv \\
 &\quad \left. + \int_{\mathbf{R}^{2d}} \varphi(t_{n+1}) (\bar{f}_{h,R}^{n+1} - f_{h,R}^{n+1}) dx dv \right] \\
 &= \int_0^T \int_{\mathbf{R}^{2d}} f_{h,R}(t) (-\Delta_v)^s \varphi(t) dt dx dv - \int_{\mathbf{R}^{2d}} \varphi(0) f_0(x, v) dx dv \\
 &\quad + \sum_{n=1}^N \int_{\mathbf{R}^{2d}} \varphi(t_n) (\tilde{f}_{h,R}^n - \bar{f}_{h,R}^n) dx dv + \sum_{n=1}^N \int_{\mathbf{R}^{2d}} \varphi(t_n) (\bar{f}_{h,R}^n - f_{h,R}^n) dx dv.
 \end{aligned}$$

In the above we also used the fact that  $\varphi$  is compactly supported in  $(-T, T)$  so that  $\varphi(t_N) = 0$ . Since  $P_{h,R}^n(dx dv dx' dv')$  is the optimal coupling in the definition of



$\mathcal{W}_h(\bar{f}_{h,R}^n, f_{h,R}^n)$ , it is easy to see that

$$\begin{aligned}
 (5.9) \quad & \int_{\mathbf{R}^{2d}} [f_{h,R}^n - \bar{f}_{h,R}^n] \varphi(t_n) dx dv \\
 &= \int_{\mathbf{R}^{2d}} f_{h,R}^n \varphi(t_n, x', v') dx' dv' - \int_{\mathbf{R}^{2d}} \bar{f}_{h,R}^n(x, v) \varphi(t_n, x, v) dx dv \\
 &= \int_{\mathbf{R}^{4d}} [\varphi(t_n, x', v') - \varphi(t_n, x, v)] P_{h,R}^n(dx dv dx' dv') \\
 &= \int_{\mathbf{R}^{4d}} [(x' - x) \cdot \nabla_{x'} \varphi(t_n, x', v') + (v' - v) \cdot \nabla_{v'} \varphi(t_n, x', v')] P_{h,R}^n(dx dv dx' dv') \\
 &\quad + \varepsilon_n,
 \end{aligned}$$

where we have used Taylor expansion in the last equality and the error term  $\varepsilon_n$  can be bounded as

$$(5.10) \quad |\varepsilon_n| \leq \frac{1}{2} \|\nabla^2 \varphi(t_n)\|_\infty \int_{\mathbf{R}^{4d}} [|x' - x|^2 + |v' - v|^2] P_{h,R}^n(dx dv dx' dv').$$

In view of (4.3), (4.4) and (5.9), we have that

$$\begin{aligned}
 (5.11) \quad & \int_{\mathbf{R}^{2d}} [f_{h,R}^n(x, v) - \bar{f}_{h,R}^n(x, v)] \varphi(t_n, x, v) dx dv \\
 &= h \int_{\mathbf{R}^{2d}} [v \cdot \nabla_x \varphi(t_n, x, v) - \nabla_v \Psi(v) \cdot \nabla_v \varphi(t_n, x, v)] f_{h,R}^n(x, v) dx dv \\
 &\quad + \frac{h^2}{2} \int_{\mathbf{R}^{2d}} \nabla_v \Psi(v) \cdot \nabla_x \varphi(t_n, x, v) f_{h,R}^n(dx dv) + \varepsilon_n.
 \end{aligned}$$

As a result the last term on the right-hand side of (5.8) can be written as

$$\begin{aligned}
 (5.12) \quad & \sum_{n=1}^N \int_{\mathbf{R}^{2d}} [\bar{f}_{h,R}^n(x, v) - f_{h,R}^n(x, v)] \varphi(t_n, x, v) dx dv \\
 &= -h \sum_{n=1}^N \int_{\mathbf{R}^{2d}} [v \cdot \nabla_x \varphi(t_n, x, v) - \nabla_v \Psi(v) \cdot \nabla_v \varphi(t_n, x, v)] f_{h,R}^n(x, v) dx dv \\
 &\quad - \frac{h^2}{2} \sum_{n=1}^N \int_{\mathbf{R}^{2d}} \nabla_v \Psi(v) \cdot \nabla_x \varphi(t_n, x, v) f_{h,R}^n(dx dv) - \sum_{n=1}^N \varepsilon_n.
 \end{aligned}$$

Now using again the fact that  $\varphi(t_N) = 0$ , we rewrite the first term on the right side of (5.12) as follows

$$\begin{aligned}
(5.13) \quad & -h \sum_{n=1}^N \int_{\mathbf{R}^{2d}} [v \cdot \nabla_x \varphi(t_n, x, v) - \nabla_v \Psi(v) \cdot \nabla_v \varphi(t_n, x, v)] f_{h,R}^n(x, v) dx dv \\
& = - \sum_{n=1}^{N-1} \int_{t_n}^{t_{n+1}} \int_{\mathbf{R}^{2d}} [v \cdot \nabla_x \varphi(t_n, x, v) - \nabla_v \Psi(v) \cdot \nabla_v \varphi(t_n, x, v)] f_{h,R}^n(x, v) dx dv dt \\
& = - \int_0^T \int_{\mathbf{R}^{2d}} [v \cdot \nabla_x \varphi(t, x, v) - \nabla_v \Psi(v) \cdot \nabla_v \varphi(t, x, v)] f_{h,R}(x, v) dx dv dt \\
& \quad + \sum_{n=1}^{N-1} \int_{t_n}^{t_{n+1}} \int_{\mathbf{R}^{2d}} \left( [v \cdot \nabla_x \varphi(t, x, v) - \nabla_v \Psi(v) \cdot \nabla_v \varphi(t, x, v)] f_{h,R}(x, v) \right. \\
& \quad \quad \left. - [v \cdot \nabla_x \varphi(t_n, x, v) - \nabla_v \Psi(v) \cdot \nabla_v \varphi(t_n, x, v)] f_{h,R}^n(x, v) \right) dx dv dt \\
& \quad + \int_0^h \int_{\mathbf{R}^{2d}} \Phi_s(t) * f_0 [v \cdot \nabla_x \varphi(t, x, v) - \nabla_v \Psi(v) \cdot \nabla_v \varphi(t, x, v)] dx dv dt.
\end{aligned}$$

Therefore the lemma follows by combining (5.8), (5.12) and (5.13).  $\square$

**5.2. Passing to the limit.** Now we set the truncation parameter  $R = h^{-1/2}$  and define

$$(5.14) \quad f_h(t) := f_{h,h^{-1/2}}(t), t \in [0, T].$$

Our aim is to prove that  $f_h$  converges to a weak solution of (1.1). To this end, we first show that the residual term in the last lemma goes to zero when  $h \rightarrow 0$ .

**Lemma 5.2.** *Let  $f_0$  be a non-negative function such that  $f_0 \in \mathcal{P}_a^2(\mathbf{R}^{2d})$  and  $\int_{\mathbf{R}^{2d}} f_0(x, v) \Psi(v) dv dx < \infty$ . Then as  $h \rightarrow 0$ , we have that*

$$(5.15) \quad |\mathcal{R}(h, h^{-1/2})| \leq C(h^2 + h + h^s + h^{1/s}) \rightarrow 0.$$

*Proof.* The proof follows closely the proof of Lemma 5.3 of [8]. In particular, by using the same arguments there, we can first obtain the following estimates

$$\begin{aligned}
\mathcal{R}_1(h, R) & \leq CT \sup_{t \in [0, T]} \|\varphi(t)\|_\infty R^{-2s}, \\
\mathcal{R}_2(h, R) & \leq \frac{Th}{2} \sup_{t \in [0, T]} \|v \cdot \nabla_x \partial_t \varphi(t, x, v) - \nabla_v \Psi(v) \cdot \nabla_v \partial_t \varphi(t, x, v)\|_\infty \\
& \quad + \frac{Th}{2} \sup_{t \in [0, T]} \left\| (-\Delta)^s \left( v \cdot \nabla_x \varphi(t, x, v) - \nabla_v \Psi(v) \cdot \nabla_v \varphi(t, x, v) \right) \right\|_\infty, \\
\mathcal{R}_3(h, R) & \leq h \sup_{t \in [0, T]} \|v \cdot \nabla_x \varphi(t, x, v) - \nabla_v \Psi(v) \cdot \nabla_v \varphi(t, x, v)\|_\infty.
\end{aligned}$$

Notice that the supreme norms appearing in the above are finite since  $\varphi \in C_0^\infty((-T \times T) \times \mathbf{R}^{2d})$  and  $\Psi \in C^{1,1} \cap C^{2,1}(\mathbf{R}^d)$ . Next, we can bound  $\mathcal{R}_4(h, R)$  as

$$\mathcal{R}_4(h, R) \leq \frac{Th}{2} \sup_{t \in [0, T]} \|\nabla_v \Psi(v) \cdot \nabla_x \varphi(t, x, v)\|_\infty.$$

In addition, thanks to inequality (4.13) and Lemma 4.1, the error term  $\tilde{\mathcal{R}}$  can be bounded as follows

$$\begin{aligned}
\tilde{\mathcal{R}}(h, R) &\leq C \sum_{n=1}^N \mathcal{W}_h(\bar{f}_{h,R}^n, f_{h,R}^n)^2 + Ch^2 \sum_{n=1}^N \left( M_{2,v}(\bar{f}_{h,R}^n) + M_{2,v}(f_{h,R}^n) \right) \\
&\leq C(1+h^2) \sum_{n=1}^N \mathcal{W}_h(\bar{f}_{h,R}^n, f_{h,R}^n)^2 + Ch^2 M_{2,v}(f^0) \\
&\quad + C(N+1)Nh^2(h^{1/s} + hR^{2-2s}) \\
&\leq C \left( h \int_{\mathbf{R}^{2d}} \Psi(v) f_0(x, v) dx dv + T \|D^2 \Psi\|_{\infty} (h^{1/s} + hR^{2-2s}) \right) \\
&\quad + Ch^2 M_{2,v}(f^0) + C(T+1)T(h^{1/s} + hR^{2-2s}).
\end{aligned}$$

Finally, the desired estimate (5.15) follows by combining the above estimates and by setting  $R = h^{-1/2}$ .  $\square$

Now we are ready to prove the main Theorem 1.3.

*Proof of Theorem 1.3.* First, thanks to Lemma 4.3 and the assumption that  $f_0 \in L^p(\mathbf{R}^{2d})$  for some  $1 < p < \infty$ , the constructed time-interpolation  $\{f_h\}$  in (5.14) is uniformly bounded in  $L^p(\mathbf{R}^{2d} \times (0, T))$ . Therefore there exists a  $f \in L^p(\mathbf{R}^{2d} \times (0, T))$  such that  $f_h \rightharpoonup h$  in  $L^p(\mathbf{R}^{2d} \times (0, T))$ . In view of equation (5.1) of Lemma 5.1, and by using the fact that  $\partial_t \varphi + v \cdot \nabla_x \varphi - \nabla_v \Psi \cdot \nabla_v \varphi - (-\Delta_v)^s \varphi \in L^{p'}(\mathbf{R}^{2d} \times (0, T))$ , we obtain by letting  $h \rightarrow 0$  that

$$\begin{aligned}
&\int_0^T \int_{\mathbf{R}^{2d}} f [\partial_t \varphi + v \cdot \nabla_x \varphi - \nabla_v \Psi \cdot \nabla_v \varphi - (-\Delta_v)^s \varphi] dx dv dt \\
&\quad + \int_{\mathbf{R}^{2d}} f_0(x, v) \varphi(0, x, v) dx dv = 0.
\end{aligned}$$

$\square$

*Remark 5.3.* By using the similar technique as in the proof of Lemma 5.8 of [8], one can show that the weak solution  $f$  of (1.1) is indeed a probability density for every  $t \in (0, T)$ , i.e.  $\int_{\mathbf{R}^{2d}} f(t, x, v) dx dv = \int_{\mathbf{R}^{2d}} f_0(x, v) dx dv = 1$ .

## 6. POSSIBLE EXTENSIONS TO MORE COMPLEX SYSTEMS

With suitable adaptations, it should be possible, in principle, to extend the analysis of the present work to deal with more complex systems. Below we briefly discuss two such systems.

**6.1. FKFPE with external force fields.** When an external force field, which is assumed to be conservative, is present, the SDE (1.2) becomes

$$\begin{aligned}
(6.1) \quad &\frac{dX_t}{dt} = V_t, \\
&\frac{dV_t}{dt} = -\nabla U(X_t) - \nabla \Psi(V_t) + dL_t^s,
\end{aligned}$$

where  $U : \mathbf{R}^d \rightarrow \mathbf{R}$  is the external potential. The corresponding FKFPE (1.1) is then given by

$$\begin{aligned}
(6.2) \quad &\begin{cases} \partial_t f + v \cdot \nabla_x f = \operatorname{div}_v(\nabla V(x) f) + \operatorname{div}_v(\nabla \Psi(v) f) - (-\Delta_v)^s f & \text{in } \mathbf{R}^d \times \mathbf{R}^d \times (0, \infty), \\ f(x, v, 0) = f_0(x, v) & \text{in } \mathbf{R}^d \times \mathbf{R}^d. \end{cases}
\end{aligned}$$

One can view (6.1) as a dissipative (frictional and stochastic noise) perturbation of the classical Hamiltonian

$$\begin{aligned}\frac{dX_t}{dt} &= V_t, \\ \frac{dV_t}{dt} &= -\nabla U(X_t).\end{aligned}$$

Thus FKFPE (6.2) contains both conservative and dissipative effects. To construct an approximation scheme for it, instead of the minimal acceleration cost function (3.2), one would use the following minimal Hamiltonian cost function which has been introduced in [19] for the development of a variational scheme for the classical Kramers equation:

(6.3)

$$\begin{aligned}\tilde{C}_h(x, v; x', v') &:= h \inf \left\{ \int_0^h |\ddot{\xi}(t) + \nabla V(\xi(t))|^2 dt : \xi \in C^1([0, h], \mathbf{R}^d) \text{ such that} \right. \\ &\quad \left. (\xi, \dot{\xi})(0) = (x, v), (\xi, \dot{\xi})(h) = (x', v') \right\}.\end{aligned}$$

Physically, the optimal value  $C_h(x, v; x', v')$  measures the least deviation from a Hamiltonian flow that connects  $(x, v)$  and  $(x', v')$  in the time interval  $[0, h]$ .

Under the assumption that  $U \in C^2(\mathbf{R}^d)$  with  $\|\nabla^2 U\| \leq C$  and using the properties of the cost function  $\tilde{C}_h$  established in [19] we expect that the splitting scheme (4.1)-(4.2), where in (3.3) the Kantorovich optimal cost functional  $C_h$  is replaced by  $\tilde{C}_h$ , can be proved to converge to a weak solution of FKFPE (6.2).

**6.2. A multi-component FKFPE equation.** The second system is an extension of FKFPE (1.1) on the phase space  $(x, v) \in \mathbf{R}^{2d}$  to a multi-component FKFPE on the space  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbf{R}^{nd}$

(6.4)

$$\begin{cases} \partial_t f + \sum_{i=2}^n x_i \cdot \nabla_{x_{i-1}} f = \operatorname{div}_{x_n} (\nabla V(x_n) f) - (-\Delta_{x_n})^s f & \text{in } \mathbf{R}^{nd} \times (0, \infty), \\ f(x_1, \dots, x_n, 0) = f_0(x_1, \dots, x_n) & \text{in } \mathbf{R}^{nd}. \end{cases}$$

Equation (6.4) with  $n > 2$  and  $s = 1$  has been studied extensively in the mathematical literature and has found many applications in different fields. For instance, it has been used as a simplified model of a finite Markovian approximation for the generalised Langevin dynamics [36, 17] or a model of a harmonic chains of oscillators that arises in the context of non-equilibrium statistical mechanics [23, 7, 15]. It has also appeared in mathematical finance [37]. Regularity properties of solutions to equation (6.4) with  $s \in (0, 1]$  has been investigated recently [29, 14, 13].

To construct an approximation scheme for equation (6.4), instead of the minimal acceleration cost function (3.2), one would use the so-called mean squared derivative cost function

$$\bar{C}_{n,h}(x_1, x_2, \dots, x_n; y_1, y_2, \dots, y_n) := h \inf_{\xi} \int_0^h |\xi^{(n)}(t)|^2 dt,$$

where  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbf{R}^{nd}$ ,  $\mathbf{y} = (y_1, \dots, y_n) \in \mathbf{R}^{nd}$ , and the infimum is taken over all curves  $\xi \in C^n([0, h], \mathbf{R}^d)$  that satisfy the boundary conditions

$$(\xi, \dot{\xi}, \dots, \xi^{(n-1)})(0) = (x_1, x_2, \dots, x_n) \quad \text{and} \quad (\xi, \dot{\xi}, \dots, \xi^{(n-1)})(h) = (y_1, y_2, \dots, y_n).$$

Several properties including an explicit representation of the mean squared derivative cost function has been studied in [20] and a variational formulation using this cost function for equation (6.4) with and  $s = 1$  has been developed recently in [21].

Using the properties of the cost function  $\bar{C}_{n,h}$  established in [20] it should be possible, in principle, to adapt the analysis of the present paper to show that, under

suitable assumptions, the splitting scheme (4.1)-(4.2) with  $C_h$  being substituted by  $\bar{C}_{n,h}$ , converges to a weak solution of the multi-component FKFPE (6.4).

#### ACKNOWLEDGMENTS

This work was partially done during the authors' stay at Warwick Mathematics Institute. The authors thank WMI for its great academic and administrative support. M. H. Duong was also supported by ERC Starting Grant 335120.

#### REFERENCES

- [1] P. Aceves-Sanchez and L. Cesbron. Fractional diffusion limit for a fractional Vlasov-Fokker-Planck equation. *SIAM J. Math. Anal.*, 51(1):469–488, 2019.
- [2] M. Agueh. Local existence of weak solutions to kinetic models of granular media. *Arch. Ration. Mech. Anal.*, 221(2):917–959, 2016.
- [3] N. Alibaud. Entropy formulation for fractal conservation laws. *J. Evol. Equ.*, 7(1):145–175, 2007.
- [4] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics. ETH Zürich. Birkhauser, Basel, 2nd edition, 2008.
- [5] D. Applebaum. *Lévy processes and stochastic calculus*, volume 116 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2009.
- [6] S. Arnrich, A. Mielke, M. A. Peletier, G. Savaré, and M. Veneroni. Passing to the limit in a Wasserstein gradient flow: From diffusion to reaction. *Calc. Var. Partial Differential Equations*, 44(3-4):419–454, 2012.
- [7] T. Bodineau and R. Lefevere. Large deviations of lattice Hamiltonian dynamics coupled to stochastic thermostats. *J. Stat. Phys.*, 133(1):1–27, 2008.
- [8] M. Bowles and M. Agueh. Weak solutions to a fractional Fokker–Planck equation via splitting and Wasserstein gradient flow. *Applied Mathematics Letters*, 42:30 – 35, 2015.
- [9] E. A. Carlen and W. Gangbo. Solution of a model Boltzmann equation via steepest descent in the 2-Wasserstein metric. *Arch. Ration. Mech. Anal.*, 172(1):21–64, 2004.
- [10] J. A. Carrillo, R. J. McCann, and C. Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana*, 19(3):971–1018, 2003.
- [11] L. Cesbron. Anomalous diffusion limit of kinetic equations in spatially bounded domains. *Comm. Math. Phys.*, 364(1):233–286, 2018.
- [12] L. Cesbron, A. Mellet, and K. Trivisa. Anomalous transport of particles in plasma physics. *Applied Mathematics Letters*, 25(12):2344 – 2348, 2012.
- [13] Z.-Q. Chen and X. Zhang. Propagation of regularity in  $L^p$ -spaces for Kolmogorov type hypoelliptic operators. [arXiv: 1706. 02181](https://arxiv.org/abs/1706.02181), 2017.
- [14] Z.-Q. Chen and X. Zhang.  $L^p$ -maximal hypoelliptic regularity of nonlocal kinetic Fokker-Planck operators. *J. Math. Pures Appl. (9)*, 116:52–87, 2018.
- [15] F. Delarue and S. Menozzi. Density estimates for a random noise propagating through a chain of differential equations. *J. Funct. Anal.*, 259(6):1577–1630, 2010.
- [16] J. Droniou, T. Gallouet, and J. Vovelle. Global solution and smoothing effect for a non-local regularization of a hyperbolic equation. *J. Evol. Equ.*, 3(3):499–521, 2003. Dedicated to Philippe Bénilan.
- [17] M. H. Duong. Long time behaviour and particle approximation of a generalised Vlasov dynamic. *Nonlinear Anal.*, 127:1–16, 2015.
- [18] M. H. Duong, A. Lamacz, M. A. Peletier, and U. Sharma. Variational approach to coarsening of generalized gradient flows. *Calc. Var. Partial Differential Equations*, 56(4):Art. 100, 65, 2017.
- [19] M. H. Duong, M. A. Peletier, and J. Zimmer. Conservative-dissipative approximation schemes for a generalized kramers equation. *Mathematical Methods in the Applied Sciences*, 37(16):2517–2540, 2014.
- [20] M. H. Duong and H. M. Tran. Analysis of the mean squared derivative cost function. *Mathematical Methods in the Applied Sciences*, 40(14):5222–5240, 2017.
- [21] M. H. Duong and H. M. Tran. On the fundamental solution and a variational formulation for a degenerate diffusion of Kolmogorov type. *Discrete Contin. Dyn. Syst.*, 38(7):3407–3438, 2018.
- [22] B. Düring, D. Matthes, and J. P. Milišić. A gradient flow scheme for nonlinear fourth order equations. *Discrete Contin. Dyn. Syst. Ser. B*, 14(3):935–959, 2010.

- [23] J.-P. Eckmann and M. Hairer. Non-equilibrium statistical mechanics of strongly anharmonic chains of oscillators. *Comm. Math. Phys.*, 212(1):105–164, 2000.
- [24] M. Erbar. Gradient flows of the entropy for jump processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 50(3):920–945, 08 2014.
- [25] W. Gangbo and M. Westdickenberg. Optimal transport for the system of isentropic Euler equations. *Comm. Partial Differential Equations*, 34(7-9):1041–1073, 2009.
- [26] P. Hänggi, P. Talkner, and M. Borkovec. Reaction-rate theory: fifty years after Kramers. *Rev. Modern Phys.*, 62(2):251–341, 1990.
- [27] H. Holden, K. H. Karlsen, K.-A. Lie, and N. H. Risebro. *Splitting methods for partial differential equations with rough solutions*. EMS Series of Lectures in Mathematics. European Mathematical Society (EMS), Zürich, 2010. Analysis and MATLAB programs.
- [28] C. Huang. A variational principle for the Kramers equation with unbounded external forces. *J. Math. Anal. Appl.*, 250(1):333–367, 2000.
- [29] L. Huang, S. Menozzi, and E. Priola.  $L^p$  estimates for degenerate non-local Kolmogorov operators. *J. Math. Pures Appl. (9)*, 121:162–215, 2019.
- [30] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the fokker-planck equation. *SIAM Journal on Mathematical Analysis*, 29(1):1–17, 1998.
- [31] D. Kinderlehrer and A. Tudorascu. Transport via mass transportation. *Discrete Contin. Dyn. Syst. Ser. B*, 6(2):311–338, 2006.
- [32] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7:284–304, 1940.
- [33] M. Kwaśnicki. Ten equivalent definitions of the fractional Laplace operator. *Fract. Calc. Appl. Anal.*, 20(1):7–51, 2017.
- [34] L. Lafleche. Fractional fokker-planck equation with general confinement force, 2018.
- [35] E. Nelson. *Dynamical theories of Brownian motion*. Princeton University Press, Princeton, N.J., 1967.
- [36] M. Ottobre and G. A. Pavliotis. Asymptotic analysis for the generalized Langevin equation. *Nonlinearity*, 24(5):1629–1653, 2011.
- [37] A. Pascucci. *Kolmogorov Equations in Physics and in Finance*, pages 353–364. Birkhäuser Basel, Basel, 2005.
- [38] H. Risken. *The Fokker-Planck Equation: Methods of Solutions and Applications*. Springer Series in Synergetics. Springer, 2nd ed. 1989. 3rd printing edition, Sept. 1996.
- [39] E. Sandier and S. Serfaty. Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Comm. Pure Appl. Math.*, 57(12):1627–1672, 2004.
- [40] U. Stefanelli. The Brezis-Ekeland principle for doubly nonlinear equations. *SIAM Journal on Control and Optimization*, 47:1615, 2008.
- [41] J. L. Vázquez. *Nonlinear Diffusion with Fractional Laplacian Operators*, pages 271–298. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.
- [42] J. L. Vázquez. Recent progress in the theory of nonlinear diffusion with fractional Laplacian operators. *Discrete Contin. Dyn. Syst. Ser. S*, 7(4):857–885, 2014.
- [43] J. L. Vázquez. The mathematical theories of diffusion: nonlinear and fractional diffusion. In *Nonlocal and nonlinear diffusions and interactions: new methods and directions*, volume 2186 of *Lecture Notes in Math.*, pages 205–278. Springer, Cham, 2017.
- [44] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

(M. H. Duong) SCHOOL OF MATHEMATICS, UNIVERSITY OF BIRMINGHAM, BIRMINGHAM B15 2TT, UK

Email address: h.duong@bham.ac.uk

(Y. Lu) DEPARTMENT OF MATHEMATICS, DUKE UNIVERSITY, DURHAM NC 27708, USA

Email address: yulonglu@math.duke.edu