

# ANALYSIS & PDE

Volume 17

No. 2

2024

FRANCESCA ANCESCHI AND YUZHE ZHU

**ON A SPATIALLY INHOMOGENEOUS NONLINEAR  
FOKKER-PLANCK EQUATION: CAUCHY PROBLEM AND  
DIFFUSION ASYMPTOTICS**

# ON A SPATIALLY INHOMOGENEOUS NONLINEAR FOKKER–PLANCK EQUATION: CAUCHY PROBLEM AND DIFFUSION ASYMPTOTICS

FRANCESCA ANCESCHI AND YUZHE ZHU

We investigate the Cauchy problem and the diffusion asymptotics for a spatially inhomogeneous kinetic model associated to a nonlinear Fokker–Planck operator. We derive the global well-posedness result with instantaneous smoothness effect, when the initial data lies below a Maxwellian. The proof relies on the hypoelliptic analog of classical parabolic theory, as well as a positivity-spreading result based on the Harnack inequality and barrier function methods. Moreover, the scaled equation leads to the fast diffusion flow under the low field limit. The relative phi-entropy method enables us to see the connection between the overdamped dynamics of the nonlinearly coupled kinetic model and the correlated fast diffusion. The global-in-time quantitative diffusion asymptotics is then derived by combining entropic hypocoercivity, relative phi-entropy, and barrier function methods.

1. Introduction	379
2. Preliminaries	385
3. Kolmogorov–Fokker–Planck equation	387
4. Well-posedness of the nonlinear model	391
5. Diffusion asymptotics	403
Appendix A. Maximum principle	414
Appendix B. Spreading of positivity	415
Appendix C. Gaining regularity of spatial increment	417
Acknowledgements	418
References	418

## 1. Introduction

We consider the kinetic Fokker–Planck operator  $\mathcal{L}_{\text{FP}} := \nabla_v \cdot (\nabla_v + v)$  and the spatially inhomogeneous nonlinear drift-diffusion model

$$\begin{cases} (\partial_t + v \cdot \nabla_x) f(t, x, v) = \rho_f^\beta(t, x) \mathcal{L}_{\text{FP}} f(t, x, v), \\ f(0, x, v) = f_{\text{in}}(x, v), \end{cases} \quad (1-1)$$

for an unknown  $f(t, x, v) \geq 0$  with  $t \in \mathbb{R}_+$ ,  $x \in \mathbb{T}^d$  or  $\mathbb{R}^d$ ,  $v \in \mathbb{R}^d$ , where  $\mathbb{T}^d$  denotes the  $d$ -dimensional torus with unit volume, the constant  $\beta \in [0, 1]$ , and

$$\rho_f(t, x) := \int_{\mathbb{R}^d} f(t, x, v) \, dv.$$

MSC2020: 35A01, 35B40, 35Q35, 35Q84.

Keywords: nonlinear kinetic Fokker–Planck equation, well-posedness, regularity, diffusion asymptotics.

Given a constant  $\epsilon \in (0, 1)$ , the equation under the low field scaling  $t \mapsto \epsilon^2 t$ ,  $x \mapsto \epsilon x$  reads

$$\begin{cases} (\epsilon \partial_t + v \cdot \nabla_x) f_\epsilon(t, x, v) = \frac{1}{\epsilon} \rho_{f_\epsilon}^\beta(t, x) \mathcal{L}_{FP} f_\epsilon(t, x, v), \\ f_\epsilon(0, x, v) = f_{\epsilon, \text{in}}(x, v). \end{cases} \tag{1-2}$$

Our aim is to show the global well-posedness and the trend to equilibrium with smoothness a priori estimates for (1-1), and the quantitative asymptotic dynamics of (1-2) as  $\epsilon$  tends to zero.

**1A. Main results.** Let us recall that a classical solution of an evolution equation is a nonnegative function satisfying the equation pointwise everywhere and matching the initial data continuously. Unless otherwise specified, any solution we consider below is intended in the classical sense. For  $k \in \mathbb{N}$ , define  $C^k(\Omega)$  to be the set of functions having all derivatives of order less than or equal to  $k$  continuous in the domain  $\Omega$ . For  $\alpha \in (0, 1)$ , we note that  $C^\alpha(\Omega)$  is the classical Hölder space on  $\Omega$  with exponent  $\alpha$ . In addition, we define the measure  $dm := dx d\mu$ , where

$$\mu(v) := (2\pi)^{-d/2} e^{-|v|^2/2} \quad \text{and} \quad d\mu := \mu dv$$

denote the Gaussian function and the Gaussian measure, respectively. A function that takes the form of  $C\mu(v)$  for some constant  $C > 0$  is called a Maxwellian.

**Theorem 1.1.** *Let the space domain  $\Omega_x$  be equal to  $\mathbb{T}^d$  or  $\mathbb{R}^d$  and the constants  $0 < \lambda < \Lambda$  be given.*

(i) *If  $f_{\text{in}} \in C^0(\Omega_x \times \mathbb{R}^d)$  satisfies  $0 \leq f_{\text{in}} \leq \Lambda\mu$  in  $\Omega_x \times \mathbb{R}^d$ , then there exists a solution  $f$  to the Cauchy problem (1-1) such that  $0 \leq f \leq \Lambda\mu$  in  $\mathbb{R}_+ \times \Omega_x \times \mathbb{R}^d$ . Moreover, for any  $v \in (0, 1)$ ,  $k \in \mathbb{N}$ , and any compact subset  $K \subset (0, T] \times \Omega_x$ , there is some constant  $C_{T,v,k,K} > 0$  depending only on  $d, \beta, \lambda, \Lambda, T, v, k, K$ , and the initial data such that*

$$\|\mu^{-v} f\|_{C^k(K \times \mathbb{R}^d)} \leq C_{T,v,k,K}.$$

*Additionally, if  $f_{\text{in}}$  is Hölder continuous and  $\rho_{f_{\text{in}}} \geq \lambda$  in  $\Omega_x$ , then the solution that lies below any Maxwellian is unique.*

(ii) *For  $\Omega_x = \mathbb{T}^d$ , if the initial data satisfies  $\lambda\mu \leq f_{\text{in}} \leq \Lambda\mu$  in  $\mathbb{T}^d \times \mathbb{R}^d$ , then, for any  $k \in \mathbb{N}$ , there exists some constant  $c > 0$  depending only on  $d, \beta, \lambda, \Lambda$  and some constant  $C_k > 0$  depending additionally on  $k$  such that, for any  $t \geq 1$ ,*

$$\left\| \frac{f - \mu \int f_{\text{in}} dx dv}{\sqrt{\mu}} \right\|_{C^k(\mathbb{T}^d \times \mathbb{R}^d)} \leq C_k e^{-ct}.$$

*For  $\Omega_x = \mathbb{R}^d$ , if the initial data satisfies  $\lambda\mu \leq f_{\text{in}} \leq \Lambda\mu$  in  $\mathbb{R}^d \times \mathbb{R}^d$  and  $f_{\text{in}} - M_1\mu \in L^1(\mathbb{R}^d \times \mathbb{R}^d)$  for some constant  $M_1 > 0$ , then there is some constant  $C' > 0$  depending only on  $d, \beta, \lambda, \Lambda, M_1$  such that*

$$\left\| \frac{f - M_1\mu}{\sqrt{\mu}} \right\|_{L^2(\mathbb{R}^d \times \mathbb{R}^d)} \leq C'(1 + \|f_{\text{in}} - M_1\mu\|_{L^1(\mathbb{R}^d \times \mathbb{R}^d)}) t^{-d/4}.$$

**Remark 1.2.** If the general measurable initial data  $f_{\text{in}}$  satisfies  $f_{\text{in}} \leq \Lambda\mu$  and an extra locally uniform lower bound assumption (see (4-14) below for a precise description), the existence of solutions still holds in some weak sense, as pointed out in Remark 4.9 below.

In order to describe the diffusion asymptotics of (1-2), we introduce the (Bregman) distance characterized by the relative phi-entropy functional  $\mathcal{H}_\beta$ .

**Definition 1.3.** Let  $\beta \in [0, 1]$ . For any measurable functions  $h_1 \geq 0$  and  $h_2 > 0$  defined in  $\mathbb{T}^d \times \mathbb{R}^d$ , the relative phi-entropy of  $h_1$  with respect to  $h_2$  is defined by

$$\mathcal{H}_\beta(h_1 | h_2) := \int_{\mathbb{T}^d \times \mathbb{R}^d} (\varphi_\beta(h_1) - \varphi_\beta(h_2) - \varphi'_\beta(h_2)(h_1 - h_2)) \, dm,$$

where  $\varphi_\beta : \mathbb{R}_+ \rightarrow \mathbb{R}$  is defined by

$$\varphi_\beta(z) := \frac{1}{1 - \beta} (z^{2-\beta} - (2 - \beta)z + 1 - \beta)$$

for  $\beta \in [0, 1)$  and  $\varphi_1(z) := z \log z - z + 1$ .

**Theorem 1.4.** Let the constants  $\alpha_0 \in (0, 1)$  and  $0 < \lambda < \Lambda$  be given, and consider a sequence of functions  $\{f_{\epsilon, \text{in}}\}_{\epsilon \in (0, 1)} \subset C^{\alpha_0}(\mathbb{T}^d \times \mathbb{R}^d)$  satisfying  $0 \leq f_{\epsilon, \text{in}} \leq \Lambda \mu$  in  $\mathbb{T}^d \times \mathbb{R}^d$  and  $\rho_{f_{\epsilon, \text{in}}} \geq \lambda$  in  $\mathbb{T}^d$ . Let  $f_\epsilon$  be the solution to (1-2) associated with the initial data  $f_{\epsilon, \text{in}}$ .

(i) If there exists some constant  $\epsilon' \in (0, 1)$  and some function  $\rho_{\text{in}} \in C^{\alpha_0}(\mathbb{T}^d)$  valued in  $[\lambda, \Lambda]$  such that

$$\mathcal{H}_\beta(\mu^{-1} f_{\epsilon, \text{in}} | \rho_{\text{in}}) \leq \epsilon',$$

then there exist constants  $M, m > 0$  depending only on  $d, \beta, \lambda, \Lambda, \alpha_0, \|\rho_{\text{in}}\|_{C^{\alpha_0}(\mathbb{T}^d)}$ , and  $\|f_{\epsilon, \text{in}}\|_{C^{\alpha_0}(\mathbb{T}^d \times \mathbb{R}^d)}$  such that, for any  $T > 0$ ,

$$\|\mu^{-1} f_\epsilon - \rho\|_{L^\infty([0, T]; L^2(\mathbb{T}^d \times \mathbb{R}^d, dm))} \leq M e^{MT} (\epsilon + \epsilon')^m,$$

where  $\rho(t, x)$  for  $(t, x) \in \mathbb{R}_+ \times \mathbb{T}^d$  is the solution to the fast diffusion equation

$$\begin{cases} \partial_t \rho(t, x) = \nabla_x \cdot (\rho^{-\beta}(t, x) \nabla_x \rho(t, x)), \\ \rho(0, x) = \rho_{\text{in}}(x). \end{cases} \tag{1-3}$$

(ii) If we additionally assume that  $f_{\epsilon, \text{in}} \geq \lambda \mu$  in  $\mathbb{T}^d \times \mathbb{R}^d$ , then there exist some constants  $M', m' > 0$  with the same dependence as  $M$  and  $m$  such that

$$\|\mu^{-1} f_\epsilon - \rho\|_{L^\infty(\mathbb{R}_+; L^2(\mathbb{T}^d \times \mathbb{R}^d, dm))} \leq M' (\epsilon + \epsilon')^{m'}.$$

### 1B. Strategy and background.

**1B1. Cauchy problem of the nonlinear model.** The well-posedness of the nonlinear model (1-1) was first studied in [Imbert and Mouhot 2021] mixing Hölder and Sobolev spaces in the torus, and in [Liao et al. 2018] under the regime of perturbation to the global equilibrium in the whole space. We develop well-posedness with rough initial data by means of the combination of the hypoelliptic analog of the parabolic theory with a positivity-spreading result; in particular, the technique we employ allows us to drop the smallness and lower bound assumptions asserted in Theorem 1.1. In addition, the global behavior of solutions to (1-1) is derived under the assumption of upper and lower bounds on the initial data only.

When the drift-diffusion coefficient  $\rho_f^\beta$  in (1-1) is proportional to the local mass of the solution — that is when  $\beta = 1$  — (1-1) and (1-2) have the same quadratic homogeneity as the Landau equation, but simpler global bounds and conservation laws. Due to the complex structure of the Landau equation, most of the existing results for its classical solutions are about the global theory under the near Maxwellian equilibrium regime [Guo 2002; Kim et al. 2020] and about the local well-posedness associated with low regularity and nonperturbative initial data [Henderson et al. 2019; 2020a]. By contrast, the boundedness from above and from below by Maxwellians of the initial data will be preserved along time for the solutions to (1-1) and (1-2), and the lack of conservation of momentum and energy of (1-2) reduces its hydrodynamic limit to the fast diffusion flow (1-3) rather than the Navier–Stokes dynamics of the scaling limit of the Landau equation, which makes its Cauchy problem and global behavior more tractable in a very general setting.

To address the nonlinear Cauchy problem subject to only one requirement that the initial data lies below a Maxwellian, we propose a method that involves several ingredients. First, in Section 3 we carry out a preliminary study on the linear counterpart of (1-1) — that is the Cauchy problem associated to the Kolmogorov operator

$$\mathcal{L}_1 := \partial_t + v \cdot \nabla_x - \text{tr}(A(t, x, v) D_v^2 \cdot) + B(t, x, v) \cdot \nabla_v, \quad (1-4)$$

where the coefficients including the entries of the positive definite  $d \times d$  real symmetric matrix  $A$  and the  $d$ -dimensional vector  $B$  are Hölder continuous ( $B$  is not necessarily bounded over  $v \in \mathbb{R}^d$ ). Even if the well-posedness theory for the Cauchy problem associated to the linear operator (1-4) was already well developed in some sense in the existing literature (see [Anceschi and Polidoro 2020; Manfredini 1997]), the Hölder spaces (see Definition 2.3) considered in those works are different from those studied in [Imbert and Mouhot 2021; Imbert and Silvestre 2021] (see Definition 2.1), which are the ones we study. Indeed, in contrast to [Imbert and Mouhot 2021], the (Schauder-type) a priori estimates proved in the previous literature are weaker and not appropriate for bootstrap arguments proving higher regularity for nonlinear problems (see Section 4C).

Secondly, the treatment of the existence issue for (1-1) in Hölder spaces is based on a fixed-point argument, where the compactness is provided by hypoelliptic regularization results; see Section 4B. A breakthrough on such a priori estimates for spatially inhomogeneous kinetic equations with a quasilinear diffusive structure in velocity was obtained in [Golse et al. 2019] and [Henderson and Snelson 2020; Imbert and Mouhot 2021], where the authors prove the kinetic (hypoelliptic) counterparts of the De Giorgi–Nash–Moser theory and the Schauder theory for classical elliptic equations (see for instance [Gilbarg and Trudinger 2001]), respectively. One may refer to [Mouhot 2018] for a summary. Armed with the Schauder estimate developed in [Imbert and Mouhot 2021] in kinetic Hölder spaces and the bootstrap procedure developed in [Imbert and Silvestre 2022] adapted to this case, we are then able to derive instantaneous  $C^\infty$  regularization for the solutions to (1-1) in Section 4C, provided that the solution is bounded from above and bounded away from vacuum, which guarantees the ellipticity in the velocity variable for (1-1).

Thirdly, in order to remove the lower bound assumption on the initial data, in Section 4A we establish a self-generating lower bound result showing that the positivity of solutions spreads everywhere instantaneously. Its proof is based on repeated applications of the spreading of positivity forward in time (see Lemma 4.5) and the spreading for all velocities (see Lemma 4.6), as proposed in

[Henderson et al. 2020b]. On the one hand, the barrier function argument will be used in the same spirit as [Henderson et al. 2020b] to show Lemma 4.5. Indeed, a *lower (resp. upper) barrier* for a certain equation is a subsolution (resp. supersolution) of the equation which bounds its solution from below (resp. above) on the boundary; it then follows from the maximum principle that the barrier function performs as a lower (resp. upper) bound of the solution. On the other hand, combining the local Harnack inequality obtained in [Golse et al. 2019] with the construction of a Harnack chain yields Lemma 4.6. We remark that the idea of the Harnack chain was first used in [Moser 1964], and an example of its application to Kolmogorov equations can be found in [Anceschi et al. 2019]. Essentially, the spreading of positivity can be seen as a lower bound estimate of the fundamental solution, which is thus related to the result in [Henderson et al. 2019], where the authors applied a probabilistic method.

A subtle point of the lower bound result lies in the possibilities of the degeneracy of solutions as  $t \rightarrow 0^+$  or  $t \rightarrow \infty$ , which leads to two delicate issues. First, with the same difficulty as mentioned in [Henderson et al. 2020a], in order to prove the uniqueness of the Cauchy problem (1-1), the nondegeneracy of diffusion up to the initial time is required so that the a priori estimates can be still applicable. We remark that, generally speaking, deriving uniqueness of solutions to nonlinear equations in rough spaces is always a classical difficulty, and the presence of a vacuum sometimes gives rise to nonuniqueness phenomenon even for the limiting equation (1-3); see for instance [Daskalopoulos and Kenig 2007]. Under the additional assumptions of Hölder continuity and absence of vacuum on the initial data, we achieve the uniqueness by using the scaling argument and Grönwall’s lemma, since the Hölder estimate around the initial time implies that the integrand in the inequality of Grönwall’s type is improved to be integrable with respect to the time variable; see the proof of Proposition 4.11 for more details. Second, we are only able to show the convergence to equilibrium if the drift-diffusion coefficient  $\rho_f^\beta$  decays slower than  $t^{-1}$  as  $t \rightarrow \infty$  in Proposition 5.1. Therefore, an additional lower Maxwellian bound on the initial data is imposed in Theorem 1.1(ii) and Theorem 1.4(ii) to ensure the solutions will be away from the vacuum uniformly along time. It would be expected that such additional lower bound assumption could be removed, especially when  $\beta$  is small.

**1B2. Long time behavior.** The drift-diffusion operator  $\mathcal{L}_{\text{FP}}$  acts only on the velocity variable and ceases to be dissipative on its unique steady state  $\mu$ , which also ensures that the null space of  $\mathcal{L}_{\text{FP}}$  is spanned by  $\mu$  and the conservation law of mass is satisfied. Consequently, the convergence to equilibrium is to be expected. With the help of the global smoothness a priori estimates, we are able to pass from the exponential convergence to equilibrium in the  $L^2$ -framework to the uniform convergence in  $C^\infty$  in Section 5A, when the spatial domain is compact—that is the periodic box  $\mathbb{T}^d$ . Therein, the  $L^2$ -convergence is obtained by the  $L^2$ -hypocoercivity under a macro-micro (fluid-kinetic) decomposition scheme, which suggests the construction of some proper entropy (Lyapunov) functional that would provide an equivalent  $L^2$ -norm for solutions. The key ingredient is to control the macroscopic part by means of the microscopic part in view of the decomposition. This hypocoercive theory was studied in [Esposito et al. 2013; Dolbeault et al. 2015; Hérau 2018] via different approaches, while their ideas are essentially the same. In [Esposito et al. 2013], the authors intended to develop the nonlinear energy estimate in an  $L^2$ -to- $L^\infty$  framework. In [Dolbeault et al. 2015] and [Hérau 2018], the authors studied

the  $L^2$ -hypocoercivity theory in an abstract setting and in the framework of pseudodifferential calculus, respectively. In addition, if the spatial domain is  $\mathbb{R}^d$  — meaning that it is not confined to a compact region — then the convergence rate slows down to an algebraic decay, for which the hypocoercive theory was captured in [Bouin et al. 2020]. We remark that the  $L^2$ -framework allows us to avoid some difficulties from the nonlinearity of the operator  $\rho_f^\beta \mathcal{L}_{\text{FP}} f$ , in contrast with  $H^1$ -entropic hypocoercivity methods (see for instance [Villani 2009]).

**1B3. Diffusion asymptotics.** The diffusion approximation serves as a simplification of collisional kinetic equations when the mean-free path is much smaller than the typical length of observation in a long time scale. This approximation for linear Fokker–Planck models can be traced back to [Degond and Mas-Gallic 1987], where the authors applied the Hilbert expansion method. One is also able to achieve the diffusion limit for (1-2) in some weak sense by applying a similar strategy to the one given in [El Ghani and Masmoudi 2010]. However, weak convergence is sometimes not appropriate for application, as a precise description of the convergence is not given. Still, the nonlinearity of the term  $\rho_{f_\epsilon}^\beta \mathcal{L}_{\text{FP}} f_\epsilon$  in (1-2) associated with nonperturbative initial data reveals some difficulties when deriving a quantitative convergence.

In order to overcome this difficulty, in Section 5B we will rely on the phi-entropy of solutions relative to their limit to see the finite-time asymptotics on the torus. The relative entropy method, which heavily relies on the regularity of solutions to the target equation, has become an effective tool in the study of hydrodynamic limits since [Bardos et al. 1993; Yau 1991] (see also [Saint-Raymond 2009]). The method applied to the diffusion asymptotics of the kinetic Fokker–Planck equation of the type with linear diffusion can be found in [Markou 2017]. The so-called phi-entropy (relative to the global equilibrium) was used to study the convergence of certain kinds of Fokker–Planck equations; see for instance [Arnold et al. 2001; Dolbeault and Li 2018]. Finally, combining the barrier function method with a careful treatment of the regularity estimate of the target equation enables us to deal with the asymptotic dynamics for the cases associated with general Hölder continuous initial data.

**1C. Physical motivation.** The spatially inhomogeneous Fokker–Planck equation (1-1) arises from modeling the evolution of some system of a large number of interacting particles from the statistical mechanical point of view. These models appear for instance in the study of plasma physics and biological dynamics; see [Chavanis 2008; Villani 2002]. Its solution can be interpreted as the probability density of the particles lying at the position  $x$  at time  $t$  with velocity  $v$ . The scaled model (1-2) for small  $\epsilon$  describes the evolution of the particle density in the small mean-free path and long-time regime, where the nondimensional parameter  $\epsilon \in (0, 1)$  designates the ratio between the mean-free path (microscopic scale) and the typical macroscopic length. The limiting equation (1-3) characterizes its macroscopic dynamics.

From the perspective of a stochastic process  $\{(X_t, V_t) : t \geq 0\}$  driven by a Brownian motion  $\{\mathcal{B}_t : t \geq 0\}$

$$\begin{cases} dX_t = V_t dt, \\ dV_t = \rho_f^\beta(t, X_t) V_t dt + \sqrt{2\rho_f^\beta(t, X_t)} d\mathcal{B}_t, \end{cases}$$

the dual equation describing the dynamics of  $\{(X_t, V_t) : t \geq 0\}$  is given by (1-1); see the review paper [Chandrasekhar 1943]. Indeed, the nonlinear term  $\rho_f^\beta \mathcal{L}_{\text{FP}} f$  models the collisional interaction of the

particles, where the mobility of these particles is hampered by their aggregation. More precisely, the nonlinear dependence on the drift-diffusion coefficient  $\rho_f^\beta$  translates the fact that the effect of friction in the interaction is positively correlated to the local mass of particles occupying the position  $x$  at time  $t$ . Moreover, the low field scaling  $t \mapsto \epsilon^2 t$ ,  $x \mapsto \epsilon x$  of (1-1) formally implies (1-2). As  $\epsilon$  tends to zero, its spatial diffusion phenomena are characterized by (1-3).

Regarding its physical interpretation, we point out that the factor multiplying the time derivative in (1-2) takes into account the long time scale. The inverse of the factor multiplying  $\rho_{f_\epsilon}^\beta \mathcal{L}_{\text{FP}} f_\epsilon$  stands for the scaled average distance traveled by particles between each collision, and it is usually referred to as mean-free path. In the small mean-free path regime, it was noticed in [Chandrasekhar 1943] that the spatial variation occurs significantly only under the long time scale that is consistent with the particle motion. In such an overdamped process, also called a low field limit or diffusion limit, the statistics of the particle motion translates into the macroscopic behavior of the particle system.

Finally, we recall that the associated phi-entropy introduced in Definition 1.3 is also known as Tsallis entropy in the physics community, which generalizes the Boltzmann–Gibbs entropy (the phi-entropy with  $\beta = 1$ ) in nonextensive statistical mechanics [Tsallis 1988]. It gives some hints for the formulation of the correlated diffusion, where the index  $\beta$  measures the degree of nonextensivity and nonlocality of the system; see [Tsallis 2009].

**1D. Organization of the paper.** The article is organized as follows. In Section 2, we recall some basic notions related to kinetic Hölder spaces that are adapted to the Fokker–Planck equations. Section 3 is devoted to the study of the linear Fokker–Planck equation with Hölder continuous coefficients. The well-posedness result Theorem 1.1(i) is proved in Section 4. The asymptotic behaviors, including Theorem 1.1(ii) and Theorem 1.4, are proved in Section 5.

## 2. Preliminaries

This section is devoted to basic notation, including the invariant structure and the kinetic Hölder space for the equations we are concerned with. Instead of the usual parabolic scaling and translations, the invariant scaling and transformation associated with the Kolmogorov operator  $\mathcal{L}_1$  (see (1-4)) is replaced by kinetic scaling and Galilean transformations, respectively. It then turns out that the appropriate Hölder space as well as its norm should be adapted to the new scaling and transformation.

**2A. The geometry associated to Kolmogorov operators.** Let  $z := (t, x, v) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d$ . We define the *kinetic scaling*

$$S_r(t, x, v) := (r^2 t, r^3 x, r v) \quad \text{for } r > 0$$

and the *Galilean transformation*

$$(t_0, x_0, v_0) \circ (t, x, v) := (t_0 + t, x_0 + x + t v_0, v_0 + v) \quad \text{for } (t_0, x_0, v_0) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d.$$

With respect to the product  $\circ$ , we are able to define the inverse of  $z$  as  $z^{-1} := (-t, -x + t v, -v)$ . In view of this structure of scaling and transformation, it is natural to define the cylinder centered at the origin of



radius  $r > 0$  as

$$Q_r := (-r^2, 0] \times B_{r^3}(0) \times B_r(0).$$

More generally, the cylinder centered at  $z_0 = (t_0, x_0, v_0)$  with radius  $r$  is defined by

$$Q_r(z_0) := \{z_0 \circ S_r(z) : z \in Q_1\} = \{(t, x, v) : t_0 - r^2 < t \leq t_0, |x - x_0 - (t - t_0)v_0| < r^3, |v - v_0| < r\}.$$

Roughly speaking, for fixed  $z_0 \in \mathbb{R}^{1+2d}$ , the Kolmogorov operator  $\mathcal{L}_1$  is invariant under the kinetic scaling and left-invariant under the Galilean transformation. It means that if  $f$  is a solution to the equation  $\mathcal{L}_1 f = 0$  in  $Q_r(z_0)$ , then  $f(z_0 \circ S_r(\cdot))$  solves an equation of the same ellipticity class in  $Q_1$ .

In addition, the associated quasinorm  $\|\cdot\|$  is defined by

$$\|z\| := \max\{|t|^{1/2}, |x|^{1/3}, |v|\},$$

and we notice that  $\|S_r(z)\| = r\|z\|$  and  $\|z_0 \circ z\| \leq 3(\|z_0\| + \|z\|)$ . For further information on the non-Euclidean geometry associated to Kolmogorov operators, one may refer to [Anceschi and Polidoro 2020; Imbert and Silvestre 2021].

**2B. Kinetic Hölder spaces and differential operators.** The proper kinetic Hölder space and kinetic degree of basic differential operators should be adapted to the above definitions such that they are homogeneous under these transformations. Their definitions were introduced in [Imbert and Silvestre 2021] (see also [Imbert and Mouhot 2021]), and we recall them below.

Given a monomial  $m(t, x, v) = t^{k_0} x_1^{k_1} \dots x_d^{k_d} v_1^{k_{d+1}} \dots v_d^{k_{2d}}$ , we define its *kinetic degree*

$$\text{deg}_{\text{kin}}(m) = 2k_0 + 3 \sum_{j=1}^d k_j + \sum_{j=d+1}^{2d} k_j.$$

Any polynomial  $p \in \mathbb{R}[t, x, v]$  can be uniquely written as a linear combination of monomials, and its kinetic degree  $\text{deg}_{\text{kin}}(p)$  is defined by the maximal kinetic degree of the monomials appearing in  $p$ . This definition is justified by the fact that  $p(S_r(z)) = r^{\text{deg}_{\text{kin}}(p)} p(z)$ .

**Definition 2.1.** Let the constant  $\alpha > 0$  and the open subset  $\Omega \subset \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d$  be given. We say a function  $f : \Omega \rightarrow \mathbb{R}$  is  $\mathcal{C}_l^\alpha$ -continuous at a point  $z_0 \in \Omega$  if there exists a polynomial  $p_0 \in \mathbb{R}[t, x, v]$  with  $\text{deg}_{\text{kin}}(p_0) < \alpha$  and a constant  $C > 0$  such that, for any  $z \in \Omega$  with  $z_0 \circ z \in \Omega$ ,

$$|f(z_0 \circ z) - p_0(z)| \leq C \|z\|^\alpha. \tag{2-1}$$

If this property holds for any  $z_0, z$  on each compact subset of  $\Omega$ , then we say  $f \in \mathcal{C}_l^\alpha(\Omega)$ . If the constant  $C$  in (2-1) is uniformly bounded for any  $z_0, z \in \Omega$ , we define the smallest value of  $C$  as the seminorm  $[f]_{\mathcal{C}_l^\alpha(\Omega)}$  and the norm  $\|f\|_{\mathcal{C}_l^\alpha(\Omega)} := [f]_{\mathcal{C}_l^\alpha(\Omega)} + \|f\|_{\mathcal{C}_l^0(\Omega)}$ , where we additionally define  $\mathcal{C}_l^0(\Omega) := \mathcal{C}^0(\Omega)$ , the space of continuous functions on  $\Omega$ , with the norm  $\|f\|_{\mathcal{C}_l^0(\Omega)} := [f]_{\mathcal{C}_l^0(\Omega)} := \|f\|_{\mathcal{C}^0(\Omega)} = \|f\|_{L^\infty(\Omega)}$ .

**Remark 2.2.** For  $\alpha \in [0, 1)$ , this  $\mathcal{C}_l^\alpha$ -continuity is equivalent to the standard definition of  $\mathcal{C}^\alpha$ -continuity with respect to the distance  $\|\cdot\|$ . The subscript “ $l$ ” of  $\mathcal{C}_l$  stems from the definition of Hölder continuity above, which is given in terms of a left-invariant distance with respect to the group structure of  $\circ$ .

We also mention another kind of Hölder space suitable for the study of Kolmogorov operators that was first used in [Manfredini 1997].

**Definition 2.3.** Let  $\alpha \in [0, 1)$  and  $\Omega \subset \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d$  be given. The space  $C_{\text{kin}}^{2+\alpha}(\Omega)$  consists of functions  $f \in C_l^0(\Omega)$  such that  $D_v^2 f, (\partial_t + v \cdot \nabla_x) f \in C_l^\alpha(\Omega)$ , and is equipped with the norm

$$\|f\|_{C_{\text{kin}}^{2+\alpha}(\Omega)} := \|f\|_{C_l^0(\Omega)} + \|D_v^2 f\|_{C_l^\alpha(\Omega)} + \|(\partial_t + v \cdot \nabla_x) f\|_{C_l^\alpha(\Omega)}.$$

The consistency between these two definitions is given by [Imbert and Silvestre 2021, Lemma 2.7] (see also [Imbert and Mouhot 2021, Lemma 2.4]), a result that we state here.

**Lemma 2.4.** Let  $\alpha \in (0, 1)$  and  $f \in C_l^{2+\alpha}(Q_2)$ . Then there exists some constant  $C > 0$  depending only on the dimension  $d$  such that

$$\|\nabla_v f\|_{C_l^\alpha(Q_1)} \leq C \|f\|_{C_l^{1+\alpha}(Q_2)} \quad \text{and} \quad \|D_v^2 f\|_{C_l^\alpha(Q_1)} + \|(\partial_t + v \cdot \nabla_x) f\|_{C_l^\alpha(Q_1)} \leq C \|f\|_{C_l^{2+\alpha}(Q_2)}.$$

**Remark 2.5.** For  $\alpha > 2$ , one can easily check that the polynomial  $p_0$  in (2-1) has the form

$$p_0(t, x, v) = f(z_0) + (\partial_t + v_0 \cdot \nabla_x) f(z_0)t + \nabla_v f(z_0) \cdot v + \frac{1}{2} D_v^2 f(z_0) v \cdot v + \dots$$

In particular, if  $\alpha \in (2, 3)$ , the polynomial expansion is independent of the  $x$ -variable.

**Remark 2.6.** A subtle difference between  $C_l^2$  and  $C_{\text{kin}}^2$  comes from the fact that, for  $f \in C_l^2$ , we have  $D_v^2 f$  and  $(\partial_t + v \cdot \nabla_x) f$  lying in  $L^\infty$  rather than  $C^0$ .

We will also employ the following notions of weighted Hölder norms in Section 3.

**Definition 2.7.** Let  $z = (t, x, v) \in \Omega := (0, T] \times \mathbb{R}^d \times \mathbb{R}^d$  with  $T \in \mathbb{R}_+$ . For  $f \in C_l^\alpha(\Omega)$  with  $\alpha > 0$  and  $\sigma \in \mathbb{R}$ , we define

$$[f]_0^{(\sigma)} := \sup_{z \in \Omega} \iota^\sigma [f]_{C_l^0(Q_\iota(z))}, \quad [f]_\alpha^{(\sigma)} := \sup_{z \in \Omega} \iota^{\alpha+\sigma} [f]_{C_l^\alpha(Q_\iota(z))}, \quad \|f\|_\alpha^{(\sigma)} := [f]_0^{(\sigma)} + [f]_\alpha^{(\sigma)},$$

where  $\iota := \min\{1, t^{1/2}\}$  measures the distance between  $z$  and the (parabolic) boundary of  $\Omega$ .

**2C. Other notation.** Throughout the article,  $B_R$  denotes the Euclidean ball in  $\mathbb{R}^d$  centered at the origin with radius  $R > 0$ . We employ the Japanese bracket defined as  $\langle \cdot \rangle := (1 + |\cdot|^2)^{1/2}$ . By abuse of notation,  $\langle \cdot \rangle$  will also denote the velocity mean in Section 5.

Moreover, we assume  $0 < \lambda < \Lambda$ . We denote by  $C$  a *universal* constant — that is to say a constant depending only on  $\beta, d, \lambda, \Lambda, \alpha, \sigma, \alpha_0$  specified in context. Finally, we write  $X \lesssim Y$  to mean that  $X \leq CY$  for some universal constant  $C > 0$ , and  $X \lesssim_q Y$  to mean that  $X \leq C_q Y$  for some  $C_q > 0$  depending only on universal constants and the quantity  $q$ .

### 3. Kolmogorov–Fokker–Planck equation

This section is devoted to the study of the Cauchy problem associated to the operator (1-4),

$$\begin{cases} \mathcal{L}_1 f := (\partial_t + v \cdot \nabla_x) f - \text{tr}(A D_v^2 f) - B \cdot \nabla_v f = s & \text{in } (0, T] \times \mathbb{R}^d \times \mathbb{R}^d, \\ f(0, x, v) = f_{\text{in}}(x, v) & \text{in } \mathbb{R}^d \times \mathbb{R}^d, \end{cases} \quad (3-1)$$

where the  $d \times d$  symmetric matrix  $A(t, x, v)$  and the  $d$ -dimensional vector  $B(t, x, v)$  satisfy the condition

$$\begin{cases} A\xi \cdot \xi \geq \lambda|\xi|^2 & \text{for any } \xi \in \mathbb{R}^d, \\ \|A\|_{C_t^\alpha} + \|B\|_{C_t^\alpha} \leq \Lambda, \end{cases} \tag{3-2}$$

where  $\alpha \in (0, 1)$  and the norm  $\|\cdot\|_{C_t^\alpha(\Omega)}$  of matrix denotes the summation of the norm of each entry. The boundedness condition at infinity means that the solution shall be bounded, which is intended for the validity of maximum principle; see the proof of [Lemma A.1](#) below.

The aim of this section is to solve the Cauchy problem (3-1) by virtue of the weighted Hölder norm ([Definition 2.7](#)) and by means of the standard continuity method combined with Schauder-type estimates. One may refer to [[Gilbarg and Trudinger 2001](#), Subsection 6.5] for the corresponding treatment in classical elliptic theory.

Throughout this section we work with the domain  $\Omega := (0, T] \times \mathbb{R}^d \times \mathbb{R}^d$ , with  $T \in \mathbb{R}_+$ . We shed light on the fact that all of the results below can be restricted to  $(0, T] \times \mathbb{T}^d \times \mathbb{R}^d$  whenever required.

**3A. Schauder estimates.** In order to apply the continuity method, first of all one needs to prove a global a priori estimate for solutions to (3-1) with respect to the weighted Hölder norm. In the kinetic setting, we have at our disposal the interior Schauder estimates proved in [[Imbert and Mouhot 2021](#), Theorem 3.9].

**Proposition 3.1** (interior Schauder estimate). *Let the constant  $\alpha \in (0, 1)$  be given and the cylinder  $Q_{2r}(z_0)$  be a subset of  $\Omega$  with  $r \in (0, 1]$ . If  $f$  satisfies (1-4) with condition (3-2) in  $Q_{2r}(z_0)$  and  $s \in C_t^\alpha(Q_{2r}(z_0))$ , then we have*

$$r^{2+\alpha} [f]_{C_t^{2+\alpha}(Q_r(z_0))} \lesssim \|f\|_{L^\infty(Q_{2r}(z_0))} + r^{2+\alpha} [s]_{C_t^\alpha(Q_{2r}(z_0))}. \tag{3-3}$$

In particular, the right-hand side controls  $r^2 \|(\partial_t + v \cdot \nabla_x) f\|_{L^\infty(Q_r(z_0))} + r^2 \|D_v^2 f\|_{L^\infty(Q_r(z_0))}$ .

First of all, we enhance this result to a global estimate for the Cauchy problem (3-1) under a vanishing condition for the initial data.

**Proposition 3.2** (global Schauder estimate). *Let  $\Omega = (0, T] \times \mathbb{R}^d \times \mathbb{R}^d$  and the constants  $\alpha \in (0, 1)$  and  $\sigma \in (0, 2)$  be universal,  $s \in C_t^\alpha(\Omega)$  such that  $\|s\|_\alpha^{(2-\sigma)} < \infty$ , and  $f$  be a bounded solution to the Cauchy problem (3-1) under condition (3-2) in  $\Omega$ . If the initial data  $f_{\text{in}}$  equals 0, then we have*

$$\|f\|_{2+\alpha}^{(-\sigma)} \lesssim \|s\|_\alpha^{(2-\sigma)}.$$

*Proof.* In view of [Proposition 3.1](#), it suffices to deal with the estimates around the initial time. Without loss of generality, we assume  $T \leq 1$ .

Let  $z_0 = (t_0, x_0, v_0) \in \Omega$  and  $2r = t_0^{1/2}$ . Applying the interior Schauder estimate (3-3) yields

$$r^{2+\alpha} [f]_{C_t^{2+\alpha}(Q_r(z_0))} \lesssim \|f\|_{L^\infty(Q_{2r}(z_0))} + r^{2+\alpha} [s]_{C_t^\alpha(Q_{2r}(z_0))}.$$

It then follows from the arbitrariness of  $z_0$  that, for any  $\sigma < 2$  such that  $[f]_0^{(-\sigma)} < \infty$ ,

$$[f]_{2+\alpha}^{(-\sigma)} \lesssim [f]_0^{(-\sigma)} + [s]_\alpha^{(2-\sigma)}. \tag{3-4}$$

With  $\sigma \in (0, 2)$ , observing that

$$\begin{aligned} \mathcal{L}_1 \left( \frac{2}{\sigma} [s]_0^{(2-\sigma)} t^{\sigma/2} \pm f \right) &= [s]_0^{(2-\sigma)} t^{\sigma/2-1} \pm s \geq 0 \quad \text{in } \Omega, \\ \frac{2}{\sigma} [s]_0^{(2-\sigma)} t^{\sigma/2} \pm f &= 0 \quad \text{on } \{t = 0\}, \end{aligned}$$

we apply the maximum principle (Lemma A.1) to the function  $(2/\sigma)[s]_0^{(2-\sigma)}t^{\sigma/2} \pm f$  to deduce that  $\pm t^{-\sigma/2} f \leq (2/\sigma)[s]_0^{(2-\sigma)}$ , which means

$$[f]_0^{(-\sigma)} \lesssim [s]_0^{(2-\sigma)}.$$

Combining this estimate with (3-4), we get the desired result. □

**3B. Cauchy problem for the linear equation.** The goal of this subsection is to prove the well-posedness of the Cauchy problem (3-1) with Hölder continuous coefficients.

**Proposition 3.3.** *Let  $\Omega = (0, T] \times \mathbb{R}^d \times \mathbb{R}^d$  and the constants  $\alpha \in (0, 1)$  and  $\sigma \in (0, 2)$  be universal. Assume that*

$$\begin{cases} A\xi \cdot \xi \geq \lambda|\xi|^2 \quad \text{for any } \xi \in \mathbb{R}^d, \\ \|A\|_{C^\alpha(\Omega)} + \|\langle v \rangle^{-1} B\|_{C^\alpha(\Omega)} \leq \Lambda. \end{cases} \tag{3-5}$$

*Then, for any  $s \in C^\alpha_1(\Omega)$  such that  $\|s\|_\alpha^{(2-\sigma)} < \infty$  and  $f_{\text{in}} \in C^0(\mathbb{R}^d \times \mathbb{R}^d)$ , there exists a unique bounded solution  $f \in C^{2+\alpha}_1(\Omega)$  to the Cauchy problem (3-1).*

**Remark 3.4.** In contrast with (3-2), condition (3-5) is weaker, which allows the coefficients of  $B$  to not necessarily be bounded globally. This fact will be applied to the Ornstein–Uhlenbeck operator  $\mathcal{L}_{\text{OU}} = (\nabla_v - v) \cdot \nabla_v$  in Section 4B.

The simplest possible setting of (3-1) under condition (3-5) is recovered by choosing  $A = I$  and  $B = 0$ , which turns out to be the classical Kolmogorov operator  $\mathcal{L}_0 := \partial_t + v \cdot \nabla_x - \Delta_v$ . This operator was first studied in [Kolmogoroff 1934], where its fundamental solution was calculated explicitly as

$$\Gamma(t, x, v) = \begin{cases} \left( \frac{\sqrt{3}}{2\pi t^2} \right)^d \exp \left( -\frac{3|x + \frac{1}{2}tv|^2}{t^3} - \frac{|v|^2}{4t} \right) & \text{for } t > 0, \\ 0 & \text{for } t \leq 0. \end{cases} \tag{3-6}$$

One can easily see that  $\Gamma$  is smooth outside of its pole (the origin). In fact, in this latter case the following result holds.

**Lemma 3.5.** *Let  $\Omega = (0, T] \times \mathbb{R}^d \times \mathbb{R}^d$  and  $\alpha \in (0, 1)$ . For any  $s \in C^\alpha_1(\Omega)$  such that  $\|s\|_\alpha^{(2-\sigma)} < \infty$ , the function*

$$f(t, x, v) = \int_{\mathbb{R}^d \times \mathbb{R}^d} \Gamma((\tau, \xi, \eta)^{-1} \circ (t, x, v)) s(\tau, \xi, \eta) \, d\tau \, d\xi \, d\eta \tag{3-7}$$

*is the unique bounded solution in  $C^{2+\alpha}_1(\Omega)$  to (3-1) with  $\mathcal{L}_1$  replaced by  $\mathcal{L}_0$  and  $f_{\text{in}} = 0$ .*

**Remark 3.6.** When the spatial domain is  $\mathbb{T}^d$ , one can apply Green’s function

$$G(t, x, v) := \sum_{n \in \mathbb{Z}^d} \Gamma(t, x + n, v),$$

which is well defined due to the decay of  $\Gamma$ .

We are now in a position to apply the standard continuity method to derive [Proposition 3.3](#).

*Proof of Proposition 3.3.* We split the proof into three steps. In the first step, we establish the case for vanishing initial data under the stronger assumption (3-2). We point out that the assumption on the coefficient  $B$  can be weakened in the second step. Finally, we deal with general continuous initial data.

Step 1. Assume  $f_{\text{in}} = 0$  and condition (3-2) holds. Let the constant  $\sigma \in (0, 2)$  be fixed and consider the Banach space  $\mathcal{Y} := (C_l^{2+\alpha}(\Omega), \|\cdot\|_{2+\alpha}^{(-\sigma)})$ . In particular, every function lying in  $\mathcal{Y}$  vanishes at  $t = 0$ .

For  $\tau \in [0, 1]$ , we define the operator  $\mathcal{L}_\tau := (1 - \tau)\mathcal{L}_0 + \tau\mathcal{L}_1$ , which can be written in the form

$$\mathcal{L}_\tau = \partial_t + v \cdot \nabla_x - \text{tr}(A_\tau D_v^2 \cdot) - \tau B \cdot \nabla_v,$$

where its coefficients  $A_\tau := (1 - \tau)I + \tau A$  and  $\tau B$  still satisfy condition (3-2) (with  $\lambda$  and  $\Lambda$  replaced by  $\min\{1, \lambda\}$  and  $\max\{1, \Lambda\}$ , respectively). For any  $w \in \mathcal{Y}$ , we have

$$\|\mathcal{L}_\tau w\|_\alpha^{(2-\sigma)} \lesssim (1 + \|A_\tau\|_\alpha^{(2)}) \|w\|_{2+\alpha}^{(-\sigma)} + \|B\|_\alpha^{(2)} \|w\|_{1+\alpha}^{(-\sigma)} \lesssim \|w\|_{2+\alpha}^{(-\sigma)}. \tag{3-8}$$

Let the set  $\mathcal{I}$  be the collection of  $\tau \in [0, 1]$  such that the Cauchy problem (3-9) is solvable for any  $s \in C_l^\alpha(\Omega)$  with  $\|s\|_\alpha^{(2-\sigma)} < \infty$ : there is a unique bounded solution  $f \in \mathcal{Y}$  satisfying

$$\begin{cases} \mathcal{L}_\tau f = s & \text{in } \Omega, \\ f(0, x, v) = 0 & \text{in } \mathbb{R}^d \times \mathbb{R}^d. \end{cases} \tag{3-9}$$

By [Lemma 3.5](#), we see that  $0 \in \mathcal{I}$ ; in particular,  $\mathcal{I}$  is not empty.

It now suffices to show that  $1 \in \mathcal{I}$ . Pick  $\tau_0 \in \mathcal{I}$ . Then the global Schauder estimate provided by [Proposition 3.2](#) implies that, for any  $s \in C_l^\alpha(\Omega)$  with  $\|s\|_\alpha^{(2-\sigma)} < \infty$ ,  $f = \mathcal{L}_{\tau_0}^{-1} s$  satisfies

$$\|\mathcal{L}_{\tau_0}^{-1} s\|_{2+\alpha}^{(-\sigma)} \lesssim \|s\|_\alpha^{(2-\sigma)}. \tag{3-10}$$

For any  $w \in \mathcal{Y}$ , since  $\tau_0 \in \mathcal{I}$  and (3-8) holds, the following Cauchy problem is solvable for any  $s \in C_l^\alpha(\Omega)$  with  $\|s\|_\alpha^{(2-\sigma)} < \infty$ :

$$\begin{cases} \mathcal{L}_{\tau_0} f = s + (\tau - \tau_0)(\mathcal{L}_0 - \mathcal{L}_1)w & \text{in } \Omega, \\ f(0, x, v) = 0 & \text{in } \mathbb{R}^d \times \mathbb{R}^d. \end{cases}$$

Thus, we can define the mapping  $F : \mathcal{Y} \rightarrow \mathcal{Y}$  by setting  $F(w) = f$ . Armed with (3-10) and (3-8), there exists a universal constant  $C > 0$  such that, for any  $u, w \in \mathcal{Y}$ ,

$$\|F(u) - F(w)\|_{2+\alpha}^{(-\sigma)} \leq C|\tau - \tau_0| \|(\mathcal{L}_0 - \mathcal{L}_1)(u - w)\|_\alpha^{(2-\sigma)} \leq C|\tau - \tau_0| \|u - w\|_{2+\alpha}^{(-\sigma)}.$$

Hence  $F$  is a contraction mapping, provided that  $|\tau - \tau_0| \leq \delta := (2C)^{-1}$ . Then,  $F$  gives a unique fixed point  $f \in \mathcal{Y}$ , which is the unique bounded solution to the Cauchy problem (3-9) in  $\mathcal{Y}$ . By dividing the interval  $[0, 1]$  into subintervals of length less than  $\delta$ , we conclude that  $1 \in \mathcal{I}$ .

**Step 2.** If  $f_{\text{in}} = 0$  and condition (3-5) holds, we approximate the coefficient  $B$  by  $B_n := B_{\varrho_n}$ , where  $\varrho_n(v) := \varrho_1(v/n)$  for  $v \in \mathbb{R}^d$ ,  $n \in \mathbb{N}_+$ , and  $\varrho_1 \in C_c^\infty(B_2)$  is valued in  $[0, 1]$  such that  $\varrho_1 \equiv 1$  in  $B_1$ . Then, for each  $n \in \mathbb{N}_+$ , the result obtained in the previous step provides a bounded solution  $f_n$  to (3-1) with  $B$  replaced by  $B_n$ . Indeed, applying the maximum principle (Lemma A.1) to the function  $\pm f - e^t \sup_\Omega |s|$  implies  $\sup_\Omega |f_n| \leq e^T \sup_\Omega |s|$ . Thanks to the interior Schauder estimate (Proposition 3.1), for any compact subset  $K \subset \Omega$ , we have that  $\{f_n\}_{n \geq N}$  is precompact in  $C_{\text{kin}}^2(K)$ , provided that  $N$  (depending on  $K$ ) is large enough. Sending  $n \rightarrow \infty$  in the equation satisfied by  $f_n$  yields that the limit function  $f \in C_l^{2+\alpha}(\Omega)$  is a bounded solution to (3-1), which satisfies  $\sup_\Omega |f| \leq e^T \sup_\Omega |s|$ .

**Step 3.** For general  $f_{\text{in}} \in C^0(\mathbb{R}^d \times \mathbb{R}^d)$ , we approximate  $f_{\text{in}}$  uniformly as  $\varepsilon \rightarrow 0$  by a sequence of smooth functions  $\{f_{\text{in}}^\varepsilon\}$  on  $\mathbb{R}^d \times \mathbb{R}^d$ . Thus, the function  $f - f_{\text{in}}^\varepsilon$  is a solution to (3-1), with the source term equal to  $s - \mathcal{L}_1 f_{\text{in}}^\varepsilon$ , and associated with the vanishing initial data. The procedure presented in the previous steps ensures a unique bounded solution  $f^\varepsilon$  to (3-1) for each  $f_{\text{in}}^\varepsilon$ .

The uniform convergence of  $\{f_{\text{in}}^\varepsilon\}$  and the maximum principle (Lemma A.1) implies the uniform convergence of  $\{f^\varepsilon\}$ . We may denote its limit by  $f \in C^0(\bar{\Omega})$ , which satisfies  $f(0, x, v) = f_{\text{in}}(x, v)$  on  $\mathbb{R}^d \times \mathbb{R}^d$ . The interior Schauder estimate again implies that  $\{f^\varepsilon\}$  is precompact in  $C_{\text{kin}}^2(K)$  for any compact subset  $K \subset \Omega$ ; then sending  $\varepsilon \rightarrow 0$  gives the solution  $f \in C_l^{2+\alpha}(\Omega)$  to (3-1). Its uniqueness is again given by the maximum principle. This concludes the proof.  $\square$

#### 4. Well-posedness of the nonlinear model

This section is devoted to the proof of Theorem 1.1(i), including a self-generating lower bound given in Section 4A, the existence and uniqueness given in Section 4B, and a smoothness a priori estimate given in Section 4C.

First, we recast the Cauchy problem (1-1) in terms of  $g(t, x, v) := \mu(v)^{-1/2} f(t, x, v)$ , an unknown function, with  $g_{\text{in}}(x, v) := \mu(v)^{-1/2} f_{\text{in}}(x, v)$  as follows:

$$\begin{cases} (\partial_t + v \cdot \nabla_x)g = \mathcal{R}[g]\mathcal{U}[g], \\ g(0, x, v) = g_{\text{in}}(x, v), \end{cases} \tag{4-1}$$

where  $\mathcal{R}[g]$  and  $\mathcal{U}[g]$  on the right-hand side are defined by

$$\mathcal{R}[g] := \left( \int_{\mathbb{R}^d} g \mu^{1/2} \, dv \right)^\beta \quad \text{and} \quad \mathcal{U}[g] := \mu^{-1/2} \nabla_v (\mu \nabla_v (\mu^{-1/2} g)) = \Delta_v g + \left( \frac{1}{2}d - \frac{1}{4}|v|^2 \right) g.$$

The main advantage of this formulation is that it allows us to get rid of the first-order term in  $v$ , and the zeroth-order term is bounded, since  $g$  is bounded from above by a Maxwellian.

For convenience, we are also concerned with the substitution  $h(t, x, v) := \mu(v)^{-1} f(t, x, v)$  and the Ornstein–Uhlenbeck operator  $\mathcal{L}_{\text{OU}} := (\nabla_v - v) \cdot \nabla_v$ . Equation (1-1) is then equivalent to

$$(\partial_t + v \cdot \nabla_x)h(t, x, v) = \mathcal{B}_h(t, x) \mathcal{L}_{\text{OU}} h(t, x, v), \quad \mathcal{B}_h(t, x) := \left( \int_{\mathbb{R}^d} h(t, x, v) \, d\mu \right)^\beta. \tag{4-2}$$

In contrast with (1-1), the zeroth-order term disappears. Let us begin by exhibiting the global bounds of solutions to (4-2) in  $(0, T) \times \mathbb{T}^d \times \mathbb{R}^d$ , which is a variant of [Imbert and Mouhot 2021, Lemma 4.1].

**Lemma 4.1** (global bounds). *Let  $a(t, x) \in L^\infty((0, T) \times \mathbb{T}^d)$  be nonnegative. Assume that, in the sense of distributions,  $h(t, x, v) \in L^\infty((0, T); H^1(\mathbb{T}^d \times \mathbb{R}^d, dm))$  satisfies  $(\partial_t + v \cdot \nabla_x)h = a \mathcal{L}_{OU} h$  in  $(0, T) \times \mathbb{T}^d \times \mathbb{R}^d$ . If  $h(0, \cdot, \cdot) \leq \Lambda$  in  $\mathbb{T}^d \times \mathbb{R}^d$ , then  $h \leq \Lambda$  in  $(0, T) \times \mathbb{T}^d \times \mathbb{R}^d$ ; if  $h(0, \cdot, \cdot) \geq \lambda$  in  $\mathbb{T}^d \times \mathbb{R}^d$ , then  $h \geq \lambda$  in  $(0, T) \times \mathbb{T}^d \times \mathbb{R}^d$ .*

*Proof.* Integrating the equation  $(\partial_t + v \cdot \nabla_x)(h - \Lambda) = a \mathcal{L}_{OU}(h - \Lambda)$  against  $(h - \Lambda)_+$  yields

$$\frac{1}{2} \int_{\mathbb{T}^d \times \mathbb{R}^d} [(h(t, \cdot, \cdot) - \Lambda)_+^2 - (h(0, \cdot, \cdot) - \Lambda)_+^2] dm = - \int_{[0, t] \times \mathbb{T}^d \times \mathbb{R}^d} a |\nabla_v (h - \Lambda)_+|^2 dt dm \leq 0$$

for any  $t \in [0, T]$ . This means that the upper bound is preserved along time. Similarly, the lower bound can be obtained by integrating the equation  $(\partial_t + v \cdot \nabla_x)(\lambda - h) = a \mathcal{L}_{OU}(\lambda - h)$  against  $(\lambda - h)_+$ .  $\square$

In particular, the above result preserving global bounds holds for solutions to (4-2) and (5-1) in  $(0, T) \times \mathbb{T}^d \times \mathbb{R}^d$ . We will also apply such result to the substitution  $g = \mu^{1/2}h$  appearing in Section 4B. Unless otherwise specified, throughout this section we set the domain  $\Omega := (0, T] \times \mathbb{T}^d \times \mathbb{R}^d$  with  $T \in \mathbb{R}_+$ . Nevertheless, as specified in Remark 4.4, Corollary 4.10, and Proposition 4.11 below, the results of this section also hold if the spatial domain is  $\mathbb{R}^d$ .

**4A. Self-generating lower bound.** Throughout this subsection, we assume that the bounded solution  $h$  of (1-1) lies below the universal constant  $\Lambda$ , which is guaranteed by Lemma 4.1 if the initial data lies below  $\Lambda$ . The aim of this subsection is to show the following positivity-spreading result. We remark that this proposition only relies on the mixing structure of the classical parabolic-type maximum principle and the transport operator, but not on the structure of the mass conservation.

**Proposition 4.2** (lower bound). *Let  $\delta > 0$ ,  $\underline{T} \in (0, T)$ , and  $h$  be a bounded solution to (4-2) in  $\Omega$  satisfying*

$$h(0, x, v) \geq \delta \mathbb{1}_{\{|x-x_0|<r, |v-v_0|<r\}} \tag{4-3}$$

for some  $(x_0, v_0) \in \mathbb{T}^d \times \mathbb{R}^d$ . Then, there exist two positive continuous functions  $\eta_1(t)$  and  $\eta_2(t)$  on  $(0, T]$  depending only on universal constants,  $T$ ,  $\delta$ ,  $r$ , and  $v_0$  such that, for any  $(t, x, v) \in \Omega$ ,

$$h(t, x, v) \geq \eta_1(t)e^{-\eta_2(t)|v|^2}. \tag{4-4}$$

**Remark 4.3.** In particular, the functions  $\eta_1(t)$  and  $\eta_2(t)$  are positive and bounded on any compact subset of  $(0, T]$ , but  $\eta_1$  might degenerate to zero and  $\eta_2$  may go to infinity as  $t$  tends to zero or infinity.

**Remark 4.4.** If one is concerned with the problem in the whole space — that is  $\Omega = (0, T] \times \mathbb{R}^d \times \mathbb{R}^d$  — we can proceed along the same lines as the proof in Appendix B to see that (4-3) implies the lower bound

$$h(t, x, v) \geq \eta_1(t, x)^{-1} e^{-\eta_2(t, x)|v|^4}, \tag{4-5}$$

where the functions  $\eta_1(t, x)$  and  $\eta_2(t, x)$  on  $(0, T] \times \mathbb{R}^d$  are positive, continuous and only depend on universal constants,  $T$ ,  $\delta$ ,  $r$ , and  $v_0$ . Compared with (4-4),  $\eta_1(t, x)$  and  $\eta_2(t, x)$  lose the uniformity in  $x$  as  $\mathbb{R}^d$  is not compact (see Step 3 of the proof of the proposition in Appendix B). In addition, the exponential tail with respect to  $v$  cannot be improved to a Gaussian type, since there is no uniform-in- $x$  lower bound on the local mass  $\int h d\mu$  such that Step 4 in Appendix B fails.

We note that the proof of the proposition is composed mainly of two lemmas. On the one hand, [Lemma 4.5](#) extends the lower bounds forward a short time from a neighborhood of any given point in  $\mathbb{T}^d \times \mathbb{R}^d$  and at any given time. On the other hand, [Lemma 4.6](#) is used to spread the lower bound for all velocities. The spreading of the lower bound in space is given by selecting the proper velocity to transport the positivity which is guaranteed by [Lemma 4.5](#). By applying these lemmas repeatedly, as proposed in [[Henderson et al. 2020b](#)], we are able to control the solution from below for any finite time. We postpone the full proof of [Proposition 4.2](#), obtained by combining these two lemmas, until [Appendix B](#).

**Lemma 4.5** (lower bound forward in time). *Let  $\delta, \tau, r \in (0, 1]$  and  $h$  be a bounded solution to (4-2) in  $\Omega$  with*

$$h(0, x, v) \geq \delta \mathbb{1}_{\{|x-x_0|<r, |v-v_0|<r/\tau\}}$$

for some  $(x_0, v_0) \in \mathbb{T}^d \times \mathbb{R}^d$ . Then there exists some universal constant  $c_0 > 0$  such that

$$h \geq \frac{1}{8} \delta \mathbb{1}_{\mathcal{P}}, \quad \mathcal{P} := \left\{ t \leq \min\{T, \tau, c_0(\tau r^{-1})^{-2} \langle v_0 \rangle^{-2}\}, |x - x_0 - tv| < \frac{1}{2}r, |v - v_0| < \frac{1}{2}r\tau^{-1} \right\}.$$

*Proof.* Let us consider the barrier function

$$\underline{h}(t, x, v) := -C_0 t + \frac{1}{2} \delta (1 - r^{-2} |x - x_0 - tv|^2 - \tau^2 r^{-2} |v - v_0|^2),$$

with the constant  $C_0 > 0$  to be determined. The region  $\mathcal{Q} := \{t \leq \min\{T, \tau\}, |x - x_0 - tv|^2 + \tau^2 |v - v_0|^2 < r^2\}$  contains  $\mathcal{P}$ . A direct computation yields

$$|\mathcal{L}_{OU} \underline{h}| \leq |\Delta_v \underline{h}| + |v \cdot \nabla_v \underline{h}| \lesssim \delta \langle \tau r^{-1} \rangle^2 \langle v_0 \rangle^2 \quad \text{in } \mathcal{Q}.$$

By choosing  $C_0 := (1/(8c_0))\delta \langle \tau r^{-1} \rangle^2 \langle v_0 \rangle^2$  for some (small) universal constant  $c_0 > 0$ , we have

$$(\partial_t + v \cdot \nabla_x - \mathcal{R}_h \mathcal{L}_{OU}) \underline{h} \leq -C_0 + \Lambda^\beta |\mathcal{L}_{OU} \underline{h}| < 0 \quad \text{in } \mathcal{Q}. \tag{4-6}$$

In addition,  $\underline{h}(t, x, v) \geq \frac{1}{8} \delta$  in  $\{t \leq c_0(\tau r^{-1})^{-2} \langle v_0 \rangle^{-2}, |x - x_0 - tv|^2 + \tau^2 |v - v_0|^2 < \frac{1}{2}r^2\} \supset \mathcal{P}$ .

Applying the classical maximum principle to  $\underline{h} - h$  in  $\mathcal{Q}$  after observing that  $\underline{h} - h \leq 0$  on the parabolic boundary  $\{t = 0\} \cap \mathcal{Q}$  and  $\{t \leq \min\{T, \tau\}, |x - x_0 - tv|^2 + \tau^2 |v - v_0|^2 = r^2\}$  yields the result.  $\square$

The spreading of lower bound to all velocities relies on the construction of a Harnack chain through iterative application of the local Harnack inequality [[Golse et al. 2019](#), Theorem 1.6]. Although some coefficients of (4-2) are unbounded globally over  $v \in \mathbb{R}^d$ , we remark that their local boundedness is sufficient for us to achieve the result through a careful study on the rescaling during the construction of the Harnack chain.

**Lemma 4.6** (lower bound for all velocities). *Let  $\delta > 0, T, R \in (0, 1], \underline{T} \in (0, T)$ , and  $h$  be a bounded solution to (4-2) in  $\Omega$  such that, for any  $t \in [0, T]$ ,*

$$h(t, x, v) \geq \delta \mathbb{1}_{\{|x-x_0-tv_0|<R, |v-v_0|<R\}} \tag{4-7}$$

for some  $(x_0, v_0) \in \mathbb{T}^d \times \mathbb{R}^d$ . Then there exists some (large) constant  $\underline{C} > 0$  depending only on universal constants,  $\underline{T}, \delta, R$ , and  $v_0$  such that, for any  $t \in [\underline{T}, T]$ , we have

$$h(t, x, v) \geq \underline{C}^{-1} e^{-\underline{C}|v|^4} \mathbb{1}_{\{|x-x_0-tv_0|<R/2\}}. \tag{4-8}$$



*Proof.* For any  $z := (t, x, v) \in \{t \in [\underline{T}, T], |x - x_0 - tv_0| < \frac{1}{2}R, v \in \mathbb{R}^d\}$ , we will construct a finite sequence of points to reach  $z$  from the region  $\{t \leq T, |x - x_0 - tv_0| < R, |v - v_0| < R\}$ , where the solution is positive by assumption. In particular,  $x$  does not exit this region. The nonlocality of the coefficient  $\mathcal{R}_h$ , with assumption (4-7), implies the nondegeneracy of the diffusion in velocity so that the positivity of the solution  $h$  propagates over  $v \in \mathbb{R}^d$  in a localized space region.

Step 1. Iterate the Harnack inequality. For  $i \in \{1, 2, \dots, N + 1\}$  with  $N \in \mathbb{N}$ , we define  $z_{N+1} := z$  and  $z_i := (t_i, x_i, v_i)$  by the relation

$$z_i = z_{i+1} \circ S_r \left( -\tau_1, 0, -\tau_2 \frac{v - v_0}{|v - v_0|} \right),$$

where the parameters  $N, r, \tau_1, \tau_2 > 0$  will be determined next. Consider the function for  $\tilde{z} := (\tilde{t}, \tilde{x}, \tilde{v}) \in Q_1$ :

$$h_i(\tilde{z}) := h(z_i \circ S_r(\tilde{z})) = h(t_i + r^2\tilde{t}, x_i + r^3\tilde{x} + r^2\tilde{t}v_i, v_i + r\tilde{v}).$$

We observe that, if the following is true for any  $\tilde{z} \in Q_1$ :

$$t_{i+1} + r^2\tilde{t} \in [0, T], \quad Nr\tau_2 \leq |v - v_0|, \tag{4-9}$$

$$|x_{i+1} + r^3\tilde{x} + r^2\tilde{t}v_{i+1} - x_0 - (t_{i+1} + r^2\tilde{t})v_0| < R, \tag{4-10}$$

then, for  $1 \leq i \leq N$ , the function  $h_{i+1}(\tilde{z})$  satisfies the equation

$$(\partial_{\tilde{t}} + \tilde{v} \cdot \nabla_{\tilde{x}})h_i = \mathcal{R}_h(z_i \circ S_r(\tilde{z}))(\Delta_{\tilde{v}}h_i - r(v_i + r\tilde{v}) \cdot \nabla_{\tilde{v}}h_i) \quad \text{in } Q_1,$$

where the coefficients satisfy

$$\delta^\beta R^{d\beta} \lesssim \mathcal{R}_h \lesssim 1 \quad \text{and} \quad |r(v_i + r\tilde{v})| \leq r(1 + |v_0| + |v - v_0|) \leq 1,$$

provided that  $r \leq (1 + |v_0| + |v - v_0|)^{-1}$ . Applying the Harnack inequality [Golse et al. 2019, Theorem 1.6] to  $h_i$  implies that there exist constants  $c_0, \tau_1 \in (0, 1)$  depending only on universal constants,  $\delta$ , and  $R$  such that, for any  $\tau_2 \in [0, 1 - \tau_1]$  and  $1 \leq i \leq N$ , we have

$$h(z_{i+1}) = h_{i+1}(0, 0, 0) \geq c_0 h_{i+1} \left( -\tau_1, 0, -\tau_2 \frac{v - v_0}{|v - v_0|} \right) = c_0 h(z_i). \tag{4-11}$$

Hence it remains to determine the chain  $\{z_i\}_{1 \leq i \leq N+1}$  such that conditions (4-9) and (4-10) hold and the point  $z_1$  stays in the region  $\{t \leq T, |x - x_0 - tv_0| < R, |v - v_0| < R\}$ .

Step 2. Determine the Harnack chain (including  $N, r$ , and  $\tau_2$ ) from a proper starting time  $t_1$ . For  $M > 0$ , we set

$$t_1 := \max \left\{ \frac{1}{2}\underline{T}, t - \frac{1}{8}R(1 + |v_0| + |v - v_0|)^{-1} \right\} \quad \text{and} \quad r := \frac{R}{M}(1 + |v_0| + |v - v_0|)^{-2}.$$

Recalling that  $T, R \in (0, 1]$ , by choosing

$$M \geq \frac{2}{\underline{T}} + \frac{\tau_1}{1 - \tau_1} \left( 8 + \frac{2}{\underline{T}} \right),$$

we have

$$r^2 \leq \frac{1}{2}\underline{T} \quad \text{and} \quad \tau_2 := \frac{r\tau_1|v - v_0|}{t - t_1} \leq 1 - \tau_1.$$

To determine the parameter  $M > 0$ , we point out that there exists some constant  $\bar{C}$  depending only on universal constants,  $\underline{T}$ ,  $\delta$ ,  $R$ , and  $v_0$  such that  $M \leq \bar{C}$  and

$$N := \frac{t - t_1}{r^2 \tau_1} \in \mathbb{N}^+.$$

Thus,  $Nr\tau_2 = |v - v_0|$ . This setting then guarantees condition (4-9).

It also follows from the iteration relation that  $v_i = v_0$ , and, for  $1 \leq i \leq N + 1$ ,

$$t_i = t_1 + (i - 1)r^2\tau_1, \quad v_i = v_0 + (i - 1)r\tau_2 \frac{v - v_0}{|v - v_0|}, \quad x_i = x - r^2\tau_1 \sum_{j=i}^N v_{j+1}. \quad (4-12)$$

**Step 3.** Determine the starting point  $x_1$ . For any  $1 \leq i \leq N$ , we estimate the departure distance from the expression (4-12)

$$|x_{i+1} - x_1 - (t_{i+1} - t_1)v_0| = \frac{1}{2}i(i + 1)r^3\tau_1\tau_2 \leq N^2r^3\tau_1\tau_2 = (t - t_1)|v - v_0| \leq \frac{1}{8}R.$$

Therefore, for any  $x \in B_{R/2}(x_0 + tv_0)$ , there exists some  $x_1 \in B_{5R/8}(x_0 + t_1v_0)$  such that  $x_{N+1} = x$ . In this setting, for any  $1 \leq i \leq N$ , we also have

$$\begin{aligned} |x_{i+1} + r^3\tilde{x} + r^2\tilde{t}v_{i+1} - x_0 - (t_{i+1} + r^2\tilde{t})v_0| \\ \leq |x_{i+1} - x_1 - (t_{i+1} - t_1)v_0| + |x_1 - x_0 - t_1v_0| + r^2|r\tilde{x} + \tilde{t}v_{i+1} - \tilde{t}v_0| \\ \leq \frac{R}{8} + \frac{5R}{8} + r^2(1 + |v - v_0|) < \frac{3R}{4} + \frac{R^2}{M^2} < R. \end{aligned}$$

Thus, condition (4-10) ensures the inequality (4-11) is satisfied for  $1 \leq i \leq N$ , which yields

$$h(t, x, v) \geq c_0^N h(t_1, x_1, v_0) \geq \delta e^{-N \log 1/c_0}.$$

Recalling that  $c_0 \in (0, 1)$  appears in (4-11) and  $N \leq T\bar{C}^2(1 + |v_0| + |v - v_0|)^4/(\tau_1 R^2)$ , we obtain the desired result (4-8). □

**4B. Existence and uniqueness.** Let us begin by summarizing some basic a priori estimates for solutions to (4-1).

**Lemma 4.7** (Hölder estimates). *Let  $\Omega_x = \mathbb{T}^d$  or  $\mathbb{R}^d$ , and let  $g$  be a solution to (4-1) in  $[0, T] \times \Omega_x \times \mathbb{R}^d$  satisfying*

$$\mathcal{R}[g] \geq \lambda \quad \text{in } [0, T] \times \Omega_x \quad \text{and} \quad 0 \leq g_{\text{in}} \leq \Lambda \mu^{1/2} \quad \text{in } \Omega_x \times \mathbb{R}^d.$$

(i) *Let  $\underline{T} \in (0, T)$  and  $\delta \in (0, \frac{1}{2})$ . There exists some universal constant  $\alpha \in (0, 1)$  such that, for any  $Q_{2r}(z_0) \subset [\underline{T}, T] \times \Omega_x \times \mathbb{R}^d$ , we have*

$$\|g\|_{C_t^{2+\alpha}(Q_r(z_0))} \lesssim_{\underline{T}, \delta} \mu^\delta(v_0). \quad (4-13)$$

(ii) *If  $g_{\text{in}} \in C^{\alpha_0}(\Omega_x \times \mathbb{R}^d)$  with (universal)  $\alpha_0 \in (0, 1)$ , then, for any  $\delta \in (0, \frac{1}{2})$ , there exists some universal constant  $\alpha \in (0, 1)$  such that*

$$\|g\|_{C_t^\alpha([0, T] \times \Omega_x \times B_1(v_0))} \lesssim_\delta (1 + [g_{\text{in}}]_{C^{\alpha_0}(\Omega_x \times \mathbb{R}^d)}) \mu^\delta(v_0).$$

We remark that, armed with [Lemma 4.1](#), the assertions (i) and (ii) in the above lemma directly follow from [\[Imbert and Mouhot 2021, Proposition 4.4\]](#) and [\[Zhu 2021, Corollary 4.6\]](#), respectively.

**Proposition 4.8** (existence). *For any  $g_{\text{in}} \in C^0(\mathbb{T}^d \times \mathbb{R}^d)$  such that  $0 \leq g_{\text{in}} \leq \Lambda\mu^{1/2}$  in  $\mathbb{T}^d \times \mathbb{R}^d$ , there exists a (classical) solution  $g$  to the Cauchy problem (4-1) satisfying  $0 \leq g \leq \Lambda\mu^{1/2}$  in  $\Omega$ .*

**Remark 4.9.** For any given nonnegative continuous function  $g_{\text{in}}$  that is not identically zero, there is some point  $(x_0, v_0) \in \mathbb{T}^d \times \mathbb{R}^d$  and some constants  $\delta, r > 0$  such that

$$g_{\text{in}}(x, v) \geq \delta \mathbb{1}_{\{|x-x_0|<r, |v-v_0|<r\}} \quad \text{in } \mathbb{T}^d \times \mathbb{R}^d. \tag{4-14}$$

We will see that the upper bound  $g_{\text{in}} \leq \Lambda\mu^{1/2}$  and the lower bound (4-14) assumptions on the initial data  $g_{\text{in}}$  (which could be discontinuous) are sufficient to ensure the existence of a solution  $g \in C^2_{\text{kin}}(\Omega)$  in the weak sense such that, for any  $\phi \in C^\infty_c([0, T] \times \mathbb{T}^d \times \mathbb{R}^d)$ ,

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} g_{\text{in}} \phi|_{t=0} = \int_{\Omega} \left\{ -g(\partial_t + v \cdot \nabla_x) \phi + \mathcal{R}[g] \nabla_v g \cdot \nabla_v \phi - \mathcal{R}[g] \left( \frac{1}{2}d - \frac{1}{4}|v|^2 \right) g \phi \right\}. \tag{4-15}$$

As solutions become regular instantaneously, the difference between the weak solution and the classical one lies only in the continuity around the initial time.

*Proof.* Let us assume that  $g_{\text{in}}$  satisfies (4-14) for some point  $(x_0, v_0) \in \mathbb{T}^d \times \mathbb{R}^d$  and some constants  $\delta, r > 0$ . By [Proposition 4.2](#), for any solution  $g$  to (4-1) and for any  $\underline{T} \in (0, T)$ , there is some  $\lambda_* > 0$  depending only on universal constants,  $\underline{T}$ ,  $T$ ,  $\delta$ ,  $r$ , and  $v_0$  such that

$$\mathcal{R}[g](t, x) \geq \lambda_* \quad \text{in } [\underline{T}, T] \times \mathbb{T}^d. \tag{4-16}$$

**Step 1.** We first approximate the initial data  $g_{\text{in}}$  by  $g_{\text{in}}^\varepsilon := g_{\text{in}} * \varrho_\varepsilon + \varepsilon\mu^{1/2}$ , where  $\varepsilon \in (0, 1]$ ,  $\varrho_\varepsilon(x, v) := \varepsilon^{-2d} \varrho_1(x/\varepsilon, v/\varepsilon)$  with  $(x, v) \in \mathbb{T}^d \times \mathbb{R}^d$ , and  $\varrho_1 \in C^\infty_c(B_1 \times B_1)$  is a nonnegative bump function such that  $\int_{\mathbb{R}^{2d}} \varrho_1 = 1$ . Then

$$\varepsilon\mu^{1/2} \leq g_{\text{in}}^\varepsilon \leq (1 + \Lambda)\mu^{1/2} \quad \text{in } T^d \times \mathbb{R}^d.$$

Let us fix  $\varepsilon \in (0, 1]$ . In order to establish the existence of a solution to (4-1) associated with the initial data  $g_{\text{in}}^\varepsilon$ , we find a fixed point of the mapping  $F : w \mapsto g$  defined by solving the Cauchy problem

$$\begin{cases} (\partial_t + v \cdot \nabla_x)g = \mathcal{R}[w]\mathcal{U}[g] & \text{in } \Omega, \\ g(0, \cdot, \cdot) = g_{\text{in}}^\varepsilon & \text{in } \mathbb{T}^d \times \mathbb{R}^d \end{cases} \tag{4-17}$$

on the closed convex subset  $\mathcal{K}$  of the Banach space  $C^\gamma_\ell(\overline{\Omega})$ ,

$$\mathcal{K} := \{w \in C^\gamma_\ell(\overline{\Omega}) : \|w\|_{C^\gamma_\ell(\Omega)} \leq \mathcal{N}, \varepsilon\mu^{1/2} \leq w \leq (1 + \Lambda)\mu^{1/2} \text{ in } \overline{\Omega}\},$$

where the constants  $\gamma \in (0, 1)$  and  $\mathcal{N} > 0$  are to be determined. We remark that (4-17) is equivalent to

$$(\partial_t + v \cdot \nabla_x)(\mu^{-1/2}g) = \mathcal{R}[w]\mathcal{L}_{\text{OU}}(\mu^{-1/2}g).$$

By [Lemma 4.1](#) and the fact that  $\mathcal{R}[w] \geq \varepsilon$ , we have  $\varepsilon\mu^{1/2} \leq g \leq (1 + \Lambda)\mu^{1/2}$  in  $\overline{\Omega}$ . In particular, for any  $w \in \mathcal{K}$ , we have the following for the lower-order term:  $|\mathcal{R}[w](\frac{1}{2}d - \frac{1}{4}|v|^2)g| \lesssim 1$ . Thus, the global Hölder estimate [\[Zhu 2021, Corollary 4.6\]](#) implies that there exist some constants  $\gamma \in (0, 1)$  and  $\mathcal{N} > 0$  depending

only on universal constants and  $\varepsilon$  such that  $\|g\|_{C_t^{2\gamma}(\Omega)} \leq \mathcal{N}$ , which also implies that the lower-order term  $|\mathcal{R}[w](\frac{1}{2}d - \frac{1}{4}|v|^2)g|$  is bounded in  $C_t^{2\gamma}(\Omega)$ . It then follows from [Proposition 3.3](#) with the interior Schauder estimate ([Proposition 3.1](#)) that the mapping  $F : \mathcal{K} \rightarrow \mathcal{K} \cap C_t^{2\gamma}(\bar{\Omega}) \cap C_t^{2+2\gamma}(\Omega)$  is well defined. In addition, with the help of the Arzelà–Ascoli theorem, we know that  $F(\mathcal{K})$  is precompact in  $C_t^\gamma(\bar{\Omega})$ .

As far as the continuity of  $F$  is concerned, we take a sequence  $\{w_n\}$  converging to  $w_\infty$  in  $C_t^\gamma(\bar{\Omega})$ . Since  $\{F(w_n)\}$  is precompact in  $C_t^\gamma(\bar{\Omega})$ , there exists a converging subsequence whose limit is  $g_\infty \in C_t^\gamma(\bar{\Omega})$  which satisfies  $g_\infty(0, \cdot, \cdot) = g_{\text{in}}^\varepsilon$  in  $\mathbb{T}^d \times \mathbb{R}^d$ . In view of the interior Schauder estimate ([Proposition 3.1](#)),  $\{F(w_n)\}$  is precompact in  $C_{\text{kin}}^2(K)$  for any compact subset  $K \subset \Omega$  and  $g_\infty \in C_{\text{kin}}^2(\Omega) \cap C^0(\bar{\Omega})$ . Sending  $n \rightarrow \infty$  in (4-17) satisfied by  $(w, g) = (w_n, F(w_n))$ , we see that (4-17) also holds for the pair of limits  $(w, g) = (w_\infty, g_\infty)$ . Then, applying the maximum principle ([Lemma A.1](#)) to

$$\begin{cases} (\partial_t + v \cdot \nabla_x)(\mu^{-1/2}(g_\infty - F(w_\infty))) = \mathcal{R}[w_\infty] \mathcal{L}_{\text{OU}}(\mu^{-1/2}(g_\infty - F(w_\infty))) & \text{in } \Omega, \\ (g_\infty - F(w_\infty))(0, \cdot, \cdot) = 0 & \text{in } \mathbb{T}^d \times \mathbb{R}^d, \end{cases}$$

we arrive at  $g_\infty = F(w_\infty)$ .

Then, for every  $\varepsilon \in (0, 1]$ , we are allowed to apply the Schauder fixed-point theorem (see for instance [\[Gilbarg and Trudinger 2001, Corollary 11.2\]](#)) to get  $g^\varepsilon \in C_{\text{kin}}^2(\Omega) \cap C^0(\bar{\Omega})$  such that  $F(g^\varepsilon) = g^\varepsilon$ , which is a (classical) solution to (4-1) associated with the initial data  $g_{\text{in}}^\varepsilon$ .

**Step 2.** Passage to the limit. Recalling the lower bound (4-16) on the coefficient and the higher-order Hölder estimate given by [Lemma 4.7\(i\)](#), for any  $\underline{T} \in (0, T)$ , we point out that  $\{g^\varepsilon\}$  is uniformly bounded in  $C_t^{2+\alpha_*}([\underline{T}, T] \times \mathbb{T}^d \times \mathbb{R}^d)$  for some constant  $\alpha_* \in (0, 1)$  with the same dependence as  $\lambda_*$ . Hence  $g^\varepsilon$  converges uniformly to  $g$  in  $C_{\text{kin}}^2([\underline{T}, T] \times \mathbb{T}^d \times \mathbb{R}^d)$ , up to a subsequence.

Write the equation satisfied by  $g^\varepsilon$  in the weak formulation: that is, for any  $\phi \in C_c^\infty(\bar{\Omega})$ ,

$$\begin{aligned} & \int_{\mathbb{T}^d \times \mathbb{R}^d} [g^\varepsilon(T, x, v)\phi(T, x, v) - g_{\text{in}}^\varepsilon(x, v)\phi(0, x, v)] \\ &= \int_{\Omega} \{g^\varepsilon(\partial_t + v \cdot \nabla_x)\phi - \mathcal{R}[g^\varepsilon]\nabla_v g^\varepsilon \cdot \nabla_v \phi + \mathcal{R}[g^\varepsilon](\frac{1}{2}d - \frac{1}{4}|v|^2)g^\varepsilon\phi\}. \end{aligned} \quad (4-18)$$

Combining the energy estimate derived by choosing  $\phi = g^\varepsilon$  above with the upper bound of  $g^\varepsilon$  provided by [Lemma 4.1](#), we have

$$\int_{\Omega} |\mathcal{R}[g^\varepsilon]\nabla_v g^\varepsilon|^2 \lesssim \int_{\Omega} \mathcal{R}[g^\varepsilon]|\nabla_v g^\varepsilon|^2 \leq \frac{1}{2} \int_{\mathbb{T}^d \times \mathbb{R}^d} |g_{\text{in}}^\varepsilon|^2 + \int_{\Omega} \mathcal{R}[g^\varepsilon](\frac{1}{2}d - \frac{1}{4}|v|^2)|g^\varepsilon|^2 \lesssim 1.$$

Therefore, after passing to a subsequence,  $\mathcal{R}[g^\varepsilon]\nabla_v g^\varepsilon$  converges weakly in  $L^2(\Omega)$ . On account of its local uniform convergence, we know that its weak limit is  $\mathcal{R}[g]\nabla_v g$ . In addition, since  $\mu^{-1/2}g^\varepsilon$  is uniformly bounded, by their local uniform convergence, we can also derive that the sequences  $g^\varepsilon$  and  $\mathcal{R}[g^\varepsilon](\frac{1}{2}d - \frac{1}{4}|v|^2)g^\varepsilon$  converge to  $g$  and  $\mathcal{R}[g](\frac{1}{2}d - \frac{1}{4}|v|^2)g$ , respectively, weakly in  $L^2(\Omega)$ , up to a subsequence. Then, for any  $\phi \in C_c^\infty([0, T] \times \mathbb{T}^d \times \mathbb{R}^d)$ , sending  $\varepsilon \rightarrow 0$  in (4-18) gives (4-15).

Furthermore, if the initial data  $g_{\text{in}}$  is continuous, then the barrier function method shows that the continuity around the initial time depends only on the upper bound of the solution and the continuity of  $g_{\text{in}}$ ; see the derivation of the estimate (5-30) of a general type in [Section 5B](#). Indeed, by (5-30)

(with  $\epsilon = 1$ ,  $R = |v_1|$ ,  $h_\epsilon = \mu^{-1/2}g$ , and  $h_{\epsilon, \text{in}} = \mu^{-1/2}g_{\text{in}}$ ), we see that, for any fixed  $\delta \in (0, 1)$  and  $(x_1, v_1) \in \mathbb{T}^d \times \mathbb{R}^d$  and for any  $(t, x, v) \in [0, \delta/(4(1 + |v_1|))] \times B_\delta(x_1, v_1)$ ,

$$\begin{aligned} |g(t, x, v) - g_{\text{in}}(x_1, v_1)| &\lesssim \delta^{-2} \langle v_1 \rangle^2 \mu^{1/2}(v_1)t + \delta^{-2} \mu^{1/2}(v_1)(|x - x_1 - tv|^2 + |v - v_1|^2) \\ &\quad + \mu^{1/2}(v_1) \sup_{B_\delta(x_1, v_1)} |g_{\text{in}}(x, v) - g_{\text{in}}(x_1, v_1)| \\ &\lesssim \delta^{-2}(t + |x - x_1|^2 + \mu^{1/2}(v_1)|v - v_1|^2) \\ &\quad + \sup_{B_\delta(x_1, v_1)} |g_{\text{in}}(x, v) - g_{\text{in}}(x_1, v_1)|. \end{aligned} \tag{4-19}$$

This implies the continuity of the solution  $g$  around  $t = 0$  and finishes the proof. □

One may extend the above existence result to the case where the spacial domain  $\mathbb{T}^d$  is replaced by  $\mathbb{R}^d$ .

**Corollary 4.10.** *For any  $g_{\text{in}} \in C^0(\mathbb{R}^d \times \mathbb{R}^d)$  such that  $0 \leq g_{\text{in}} \leq \Lambda \mu^{1/2}$  in  $\mathbb{R}^d \times \mathbb{R}^d$ , there exists a solution  $g$  to the Cauchy problem (4-1) satisfying  $0 \leq g \leq \Lambda \mu^{1/2}$  in  $(0, T] \times \mathbb{R}^d \times \mathbb{R}^d$ . If additionally  $h_{\text{in}} = \mu^{-1/2}g_{\text{in}}$  satisfies  $h_{\text{in}} \geq \lambda$  and  $h_{\text{in}} - M_1 \in L^1(\mathbb{R}^d \times \mathbb{R}^d, dm)$  for some constant  $M_1 > 0$ , then  $h = \mu^{-1/2}g$  satisfies  $h \geq \lambda$  in  $(0, T] \times \mathbb{R}^d \times \mathbb{R}^d$  and*

$$\|h - M_1\|_{L^\infty([0, T]; L^1(\mathbb{R}^d \times \mathbb{R}^d, dm))} \leq \|h_{\text{in}} - M_1\|_{L^1(\mathbb{R}^d \times \mathbb{R}^d, dm)}. \tag{4-20}$$

*Proof.* For  $R > 1$ , we set  $g_{\text{in}}^R = g_{\text{in}} \mathbb{1}_{[-R+R^{-1}, R-R^{-1}]^d}$  for  $x \in [-R, R]^d$  with periodic extension to  $\mathbb{R}^d$ . In the light of Proposition 4.8, we take a solution  $g^R$  to (4-1) associated with the initial data  $g_{\text{in}}^R$  in  $(0, T] \times [-R, R]^d \times \mathbb{R}^d$ , where  $[-R, R]^d$  is considered as a periodic box. After extracting a subsequence, we define the function  $g := \lim_{R \rightarrow \infty} g^R$  in  $(0, T] \times \mathbb{R}^d \times \mathbb{R}^d$  pointwise; furthermore, since  $0 \leq \mu^{-1/2}g^R \leq \Lambda$  in  $(0, T] \times [-R, R]^d \times \mathbb{R}^d$ , we know that the limiting function satisfies  $0 \leq \mu^{-1/2}g \leq \Lambda$  in  $(0, T] \times \mathbb{R}^d \times \mathbb{R}^d$ . Similarly,  $\mu^{-1/2}g_{\text{in}} \geq \lambda$  in  $\mathbb{R}^d \times \mathbb{R}^d$  implies that  $\mu^{-1/2}g \geq \lambda$  in  $(0, T] \times \mathbb{R}^d \times \mathbb{R}^d$ .

Since the initial data is continuous unless it is identically zero, we assume that  $g_{\text{in}} \geq \delta \mathbb{1}_{\{|x-x_0|<r, |v-v_0|<r\}}$  for some point  $(x_0, v_0) \in \mathbb{R}^d \times \mathbb{R}^d$  and some constants  $\delta, r > 0$ . Consider  $R > |x_0| + r$ . Applying the lower bound of the solution given by (4-5) yields that, for any compact subset  $K \subset (0, T] \times \mathbb{R}^d \times \mathbb{R}^d$ , the coefficient  $\mathcal{R}[g^R]$  is greater than or equal to  $\lambda_*$ , where the constant  $\lambda_* > 0$  only depends on universal constants,  $\delta, r, v_0$ , and  $K$ . In view of the higher-order Hölder estimate given by Lemma 4.7(i), we know that  $g^R$  uniformly converges to  $g$  in  $C_{\text{kin}}^2(K)$ , up to a subsequence. Additionally, due to the estimate derived in (4-19), the limiting function  $g$  is a solution to (4-1) that matches the initial data  $g_{\text{in}}$  continuously.

As for (4-20), we notice that the function  $(h^R - M_1)_\pm$  with  $h^R := \mu^{-1/2}g^R$  satisfies

$$(\partial_t + v \cdot \nabla_x)(h^R - M_1)_\pm \leq \mathcal{B}_{h^R} \mathcal{L}_{\text{OU}}(h^R - M_1)_\pm \quad \text{in } (0, T] \times [-R, R]^d \times \mathbb{R}^d.$$

Integrating the equation against the function  $\int_{[-R, R]^d} (h^R - M_1)_\pm dx$  yields

$$\int_{[-R, R]^d \times \mathbb{R}^d} (h^R(t, \cdot, \cdot) - M_1)_\pm dm - \int_{[-R, R]^d \times \mathbb{R}^d} (h^R(0, \cdot, \cdot) - M_1)_\pm dm \leq 0.$$

Sending  $R \rightarrow \infty$ , we acquire

$$\|(h - M_1)_\pm\|_{L^\infty([0, T]; L^1(\mathbb{R}^d \times \mathbb{R}^d, dm))} \leq \|(h_{\text{in}} - M_1)_\pm\|_{L^1(\mathbb{R}^d \times \mathbb{R}^d, dm)},$$

which implies the estimate (4-20) as asserted. □

The following proposition concerned with the uniqueness of the Cauchy problem (4-1) is derived from a Grönwall-type argument. The standard scaling technique and the Hölder estimate up to the initial time given by Lemma 4.7(ii) can improve the integrability with respect to  $t$  in the energy estimate so that Grönwall’s inequality becomes admissible; see (4-25) for the precise expression. This kind of phenomena was also noticed in [Henderson et al. 2020a] (see the remarks in §1.4.2). The global energy estimate of (4-1) is not available when the spatial domain is unbounded, since there is no decay of the solution as  $|x| \rightarrow \infty$ . To work it out, we take advantage of the idea originated from the uniformly local space used in [Henderson et al. 2019; Kato 1975]. We note that such a technique is not necessary when working with the periodic box  $\mathbb{T}^d$ .

**Proposition 4.11** (uniqueness). *Let the domain  $\Omega_x$  be  $\mathbb{T}^d$  or  $\mathbb{R}^d$ , the constant  $\alpha_0 \in (0, 1)$ , and the functions  $0 \leq g_1, g_2 \lesssim \mu^{1/2}$  be two solutions to (4-1) in  $(0, T] \times \Omega_x \times \mathbb{R}^d$  associated with the same initial data  $g_{\text{in}} \in C^{\alpha_0}(\Omega_x \times \mathbb{R}^d)$  such that*

$$\int_{\mathbb{R}^d} g_{\text{in}} \mu^{1/2} \, dv \geq \lambda \quad \text{in } \Omega_x \quad \text{and} \quad 0 \leq g_{\text{in}} \lesssim \mu^{1/2} \quad \text{in } \Omega_x \times \mathbb{R}^d.$$

Then  $g_1 = g_2$  in  $[0, T] \times \Omega_x \times \mathbb{R}^d$ .

*Proof.* In view of the lower bound given by Lemma 4.5 and Proposition 4.2, we know that there is some constant  $\lambda_* \in (0, 1)$  depending only on universal constants,  $T$ , and the initial data such that

$$\int_{\mathbb{R}^d} g_i \mu^{1/2} \, dv \geq \lambda_* \quad \text{in } [0, T] \times \Omega_x, \quad i = 1, 2. \tag{4-21}$$

Therefore, we may assume  $T = \Lambda^{-1}$  with  $\Lambda > 1$ . Let us set the difference

$$\tilde{g} := \exp\left(-\frac{1}{8}|v|^2 t\right)(g_1 - g_2).$$

We have to show that  $\tilde{g}$  is identically zero.

In view of (4-1), a direct computation yields that the function  $\tilde{g}$  satisfies

$$\begin{aligned} (\partial_t + v \cdot \nabla_x) \tilde{g} + \frac{1}{8}|v|^2 \tilde{g} \\ = \exp\left(-\frac{1}{8}|v|^2 t\right) (\mathcal{R}[g_1] - \mathcal{R}[g_2]) \mathcal{U}[g_1] + \mathcal{R}[g_2] (\mathcal{U}[\tilde{g}] + \frac{1}{2} t v \cdot \nabla_v \tilde{g} + \left(\frac{1}{4} dt + \frac{1}{16}|v|^2 t^2\right) \tilde{g}), \end{aligned} \tag{4-22}$$

with the initial condition  $\tilde{g}(0, x, v) = 0$  in  $\Omega_x \times \mathbb{R}^d$ .

Let  $y \in \mathbb{R}^d$ . We introduce a cut-off function  $\phi_y(x) := \phi(x - y)$ , where  $\phi \in C_c^\infty(\mathbb{R}^d)$  is valued in  $[0, 1]$  such that  $\phi|_{B_1} \equiv 1$ ,  $\phi|_{B_2^c} \equiv 0$ , and  $|\nabla \phi| \lesssim 1$  in  $\mathbb{R}^d$ . For any  $t \in (0, T]$ , integrating (4-22) against  $\phi_y^2 \tilde{g}$  in  $\Omega_x \times \mathbb{R}^d$  and applying integration by parts yields

$$\begin{aligned} \frac{1}{2} \int_{\Omega_x \times \mathbb{R}^d} \phi_y^2 \tilde{g}^2(t) = \int_0^t \int_{\Omega_x \times \mathbb{R}^d} \left\{ (v \cdot \nabla \phi_y) \phi_y \tilde{g}^2 - \frac{1}{8}|v|^2 \phi_y^2 \tilde{g}^2 + \exp\left(-\frac{1}{8}|v|^2 t\right) (\mathcal{R}[g_1] - \mathcal{R}[g_2]) \mathcal{U}[g_1] \phi_y^2 \tilde{g} \right. \\ \left. - \mathcal{R}[g_2] (|\nabla_v(\mu^{-1/2} \tilde{g})|^2 \mu \phi_y^2 + \frac{1}{4} dt \phi_y^2 \tilde{g}^2 - \left(\frac{1}{4} dt + \frac{1}{16}|v|^2 t^2\right) \phi_y^2 \tilde{g}^2) \right\}. \end{aligned}$$

Since  $\mathcal{R}[g_2] \in [0, \Lambda]$ , for any  $t \in (0, T]$ , we have

$$\frac{1}{2} \int_{\Omega_x \times \mathbb{R}^d} \phi_y^2 \tilde{g}^2(t) \leq \int_0^t \int_{\Omega_x \times \mathbb{R}^d} \left\{ |v| |\nabla \phi_y| \phi_y \tilde{g}^2 - \frac{1}{16}|v|^2 \phi_y^2 \tilde{g}^2 + \mu^{-1/4} |\mathcal{R}[g_1] - \mathcal{R}[g_2]| |\mathcal{U}[g_1]| \phi_y^2 |\tilde{g}| \right\}.$$

Due to the elementary inequality  $|z^\beta - 1| \leq |z - 1|$  (with  $z \in \mathbb{R}^+$ ) and the lower bound estimate in (4-21), as well as the boundedness assumption on  $g_1$  and  $g_2$ , we have

$$|\mathcal{R}[g_1] - \mathcal{R}[g_2]| \leq \mathcal{R}[g_1]^{(\beta-1)/\beta} |\mathcal{R}[\tilde{g}]|^{1/\beta} \leq \frac{1}{\lambda_*} \int_{\mathbb{R}^d} |\tilde{g}(t, x, \cdot)| \mu^{1/2} \lesssim_{\lambda_*} 1 \quad \text{in } [0, 1] \times \Omega_x.$$

It then follows that, for any  $t \in (0, T]$ ,

$$\begin{aligned} \frac{1}{2} \int_{\Omega_x \times \mathbb{R}^d} \phi_y^2 \tilde{g}^2(t) &\leq \int_0^t \int_{\Omega_x \times \mathbb{R}^d} (|v| |\nabla \phi_y| \phi_y \tilde{g}^2 - \frac{1}{16} |v|^2 \phi_y^2 \tilde{g}^2) \\ &\quad + \frac{1}{\lambda_*} \int_0^t \|\mu^{-3/8} \mathcal{U}[g_1]\|_{L_{x,v}^\infty} \int_{\Omega_x \times \mathbb{R}_v^d} \phi_y^2 |\tilde{g}(t, x, v)| \mu^{1/8} \int_{\mathbb{R}_\xi^d} |\tilde{g}(t, x, \xi)| \mu^{1/2} d\xi \\ &\lesssim_{\lambda_*} \int_0^t \int_{\Omega_x \times \mathbb{R}^d} |\nabla \phi_y|^2 \tilde{g}^2 + \int_0^t \|\mu^{-3/8} \mathcal{U}[g_1]\|_{L_{x,v}^\infty} \int_{\Omega_x \times \mathbb{R}^d} \phi_y^2 \tilde{g}^2, \end{aligned} \quad (4-23)$$

where we used the Cauchy–Schwarz inequality and Hölder’s inequality in the last line. Recalling that  $\phi_y(x) = \phi(x - y) \in C_c^\infty(\mathbb{R}^d)$  and  $|\nabla \phi| \lesssim 1$  in  $\mathbb{R}^d$ , we have

$$\sup_{y \in \mathbb{R}^d} \int_{\Omega_x \times \mathbb{R}^d} |\nabla \phi_y|^2 \tilde{g}^2 \lesssim \sup_{y \in \mathbb{R}^d} \int_{\Omega_x \times \mathbb{R}^d} \phi_y^2 \tilde{g}^2.$$

By the definition of  $\mathcal{U}[g_1]$  and the upper bound  $g_1 \lesssim \mu^{1/2}$ ,

$$\|\mu^{-3/8} \mathcal{U}[g_1]\|_{L_{x,v}^\infty} \lesssim 1 + \|\mu^{-3/8} \Delta_v g_1\|_{L_{x,v}^\infty}.$$

Hence, for any  $t \in (0, T]$ , taking supremum over  $y \in \mathbb{R}^d$  in (4-23), we obtain

$$\sup_{y \in \mathbb{R}^d} \int_{\Omega_x \times \mathbb{R}^d} \phi_y^2 \tilde{g}^2(t) \lesssim_{\lambda_*} \int_0^t (1 + \|\mu^{-3/8} \Delta_v g_1\|_{L_{x,v}^\infty}) \sup_{y \in \mathbb{R}^d} \int_{\Omega_x \times \mathbb{R}^d} \phi_y^2 \tilde{g}^2. \quad (4-24)$$

Now we have to consider the pointwise estimate on  $D_v^2 g_1$ . Let  $z_0 = (t_0, x_0, v_0) \in (0, T] \times \Omega_x \times \mathbb{R}^d$  and  $2r = t_0^{1/2}$ . In view of (4-21), Lemma 4.7(ii) implies that there exists some constant  $\alpha_* \in (0, 1)$  with the same dependence as  $\lambda_*$  such that

$$\|g_1\|_{C_t^{\alpha_*}([0, T] \times \Omega_x \times B_1(v_0))} \lesssim_{\lambda_*} 1 + [g_{\text{in}}]_{C^{\alpha_0}(\Omega_x \times \mathbb{R}^d)}.$$

Then, applying the interior Schauder estimate (Proposition 3.1) and the upper bound  $g_1 \lesssim \mu^{1/2}$  yields

$$\begin{aligned} \|D_v^2 g_1\|_{L^\infty(Q_r(z_0))} &\lesssim_{\lambda_*} r^{-2} \|g_1 - g_1(z_0)\|_{L^\infty(Q_{2r}(z_0))} + r^{\alpha_*} [\mathcal{R}[g_1] (\frac{1}{2}d - \frac{1}{4}|v|^2) g_1]_{C_t^{\alpha_*}(Q_{2r}(z_0))} \\ &\lesssim_{\lambda_*} r^{-2+\alpha_*/4} \mu^{3/8}(v_0) [g_1]_{C_t^{\alpha_*}(Q_{2r}(z_0))}^{1/4} + \mu^{3/8}(v_0) [g_1]_{C_t^{\alpha_*}(Q_{2r}(z_0))}^{1/4} \\ &\lesssim_{\lambda_*} t_0^{-1+\alpha_*/8} \mu^{3/8}(v_0) (1 + [g_{\text{in}}]_{C^{\alpha_0}(\Omega_x \times \mathbb{R}^d)}^{1/4}). \end{aligned}$$

By the arbitrariness of  $z_0$ , we know that, for any  $s \in (0, T]$ ,

$$\|\mu^{-3/8} \Delta_v g_1(s)\|_{L^\infty(\Omega_x \times \mathbb{R}^d)} \lesssim_{\lambda_*} (1 + [g_{\text{in}}]_{C^{\alpha_0}(\Omega_x \times \mathbb{R}^d)}^{1/4}) s^{-1+\alpha_*/8}.$$

Dragging this estimate into (4-24) yields that for any  $t \in (0, T]$ ,

$$\sup_{y \in \mathbb{R}^d} \int_{\Omega_x \times \mathbb{R}^d} \phi_y^2 \tilde{g}^2(t) \leq C_* \int_0^t ds (1 + s^{-1+\alpha_*/8}) \sup_{y \in \mathbb{R}^d} \int_{\Omega_x \times \mathbb{R}^d} \phi_y^2 \tilde{g}^2(s), \tag{4-25}$$

where the constant  $C_* > 0$  depends only on universal constants and the initial data. The desired result is then given by Grönwall’s inequality.  $\square$

**4C. Global regularity.** The instantaneous smoothness a priori estimate in Theorem 1.1(i) is made up of the lower bound given by Proposition 4.2 and the following proposition.

**Proposition 4.12.** *Let  $\Omega_x = \mathbb{T}^d$  or  $\mathbb{R}^d$ , let  $\underline{T} \in (0, T)$ , and let the function  $g$  be a solution to (4-1) in  $(0, T) \times \mathbb{T}^d \times \mathbb{R}^d$  such that*

$$\mathcal{R}[g] \geq \lambda \quad \text{in } [\underline{T}/4, T] \times \Omega_x \quad \text{and} \quad 0 \leq g \leq \Lambda \mu^{1/2} \quad \text{in } [0, T] \times \Omega_x \times \mathbb{R}^d. \tag{4-26}$$

Then, for any  $\nu \in (0, \frac{1}{2})$  and  $k \in \mathbb{N}$ , we have

$$\|\mu^{-\nu} g\|_{C^k([\underline{T}, T] \times \Omega_x \times \mathbb{R}^d)} \leq C_{\underline{T}, \nu, k}$$

for some constant  $C_{\underline{T}, \nu, k} > 0$  depending only on universal constants,  $\underline{T}$ ,  $\nu$ , and  $k$ .

Generally speaking, if  $g$  is a solution to (4-1) in  $(0, T] \times \Omega_x \times \mathbb{R}^d$  constructed by Proposition 4.8 (with  $\Omega_x = \mathbb{T}^d$ ) or Corollary 4.10 (with  $\Omega_x = \mathbb{R}^d$ ), then the uniform positivity assumption (4-26) should be replaced by

$$\mathcal{R}[g] \geq \lambda_{t,x} \quad \text{in } (0, T] \times \Omega_x \quad \text{and} \quad 0 \leq g \leq \Lambda \mu^{1/2} \quad \text{in } \Omega_x \times \mathbb{R}^d,$$

where  $\lambda_{t,x} > 0$  may degenerate to zero as  $t \rightarrow 0$  or  $t + |x| \rightarrow \infty$ ; see Proposition 4.2 and Remark 4.4. As an immediate consequence of the above proposition, for any  $\nu \in (0, \frac{1}{2})$ ,  $k \in \mathbb{N}$ , and for any compact subset  $K \subset (0, T] \times \Omega_x$ , there exists some constant  $C_{\nu, k, K} > 0$  depending only on universal constants,  $\nu$ ,  $k$ , and  $K$  such that

$$\|\mu^{-\nu} g\|_{C^k(K \times \mathbb{R}^d)} \leq C_{\nu, k, K},$$

which is exactly the assertion in Theorem 1.1(i).

In order to show the higher regularity, we will apply the bootstrap procedure developed in [Imbert and Silvestre 2022] which was intended for the non-cut-off Boltzmann equation. The classical bootstrap iteration proceeds by differentiating the equation, using a priori estimates to the new equation to improve the regularity of solutions, and repeating the procedure. Nevertheless, since  $C_l^{2+\alpha} \not\subset C_x^1$  for any  $\alpha \in (0, 1)$  by their definitions, the hypoelliptic structure of (4-1) does not gain enough regularity in the  $x$ -variable which disables the  $x$ -differentiation at each iteration. Indeed, the Schauder-type estimate provided by Lemma 4.7(i) only shows that the solution to (4-1) belongs to  $C^{(2+\alpha)/3}$  with respect to the  $x$ -variable. In order to overcome it, we have to apply estimates to increments of the solution to recover a full derivative. From now on, for  $y \in \mathbb{R}^d$  and  $w \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d$ , we denote the spatial increment by

$$\delta_y g(z) := g(w \circ (0, y, 0)) - g(w).$$

Let us proceed with the proof of the regularity estimate.



*Proof of Proposition 4.12.* We are going to show that, for any multi-index  $k := (k_t, k_x, k_v) \in \mathbb{N} \times \mathbb{N}^d \times \mathbb{N}^d$  and  $v \in (0, \frac{1}{2})$ , there exists some constant  $\alpha_k \in (0, 1)$  depending only on  $|k|$  such that, for any  $Q_r(z_0) \subset [\frac{1}{2}T, T] \times \Omega_x \times \mathbb{R}^d$ ,

$$\|\partial_t^{k_t} \partial_x^{k_x} \partial_v^{k_v} g\|_{C_t^{2+\alpha_k}(Q_r(z_0))} \lesssim_{T,v,k} \mu^v(v_0). \tag{4-27}$$

For simplicity, we will omit the domain in estimates below, since the estimates can be always localized around the center  $z_0$ .

Step 0. The case of  $k = (0, 0, 0)$  in (4-27) is a direct consequence of Lemma 4.7(i).

Step 1. We will establish that (4-27) holds for any differential operators of the type  $\partial_x^{k_x}$ . It suffices to show that, for any  $n \in \mathbb{N}$ ,  $k_x \in \mathbb{N}^d$  with  $|k_x| = n$ ,  $v \in (0, \frac{1}{2})$ , and  $y \in B_{r^{3/4}}$ ,

$$\|\delta_y \partial_x^{k_x} g\|_{C_t^{2+\alpha_n}} \lesssim_{T,v,n} |y| \mu^v(v_0). \tag{4-28}$$

Indeed, sending  $y \rightarrow 0$  in (4-28) will complete this step.

Based on an induction on  $|k_x| = n$ , we suppose that (4-28) holds for any  $|k_x| \leq n - 1$ , which implies, for any  $k_x \in \mathbb{N}^d$  with  $|k_x| \leq n$ ,

$$\|\partial_x^{k_x} g\|_{C_t^{2+\alpha_n}} \lesssim_{T,v,n} \mu^v(v_0). \tag{4-29}$$

We remark that the induction here begins with (4-29) for  $|k_x| = 0$ , which holds due to the previous step.

Let  $q := \delta_y \partial_x^{k_x} g$  with  $|k_x| = n$ . Lemma C.2 and (4-29) gives

$$\|q\|_{C_t^{\alpha_n}} \lesssim \|\partial_x^{k_x} g\|_{C_t^{2+\alpha_n}} \|(0, y, 0)\|^2 \lesssim_{T,v,n} |y|^{2/3} \mu^v(v_0). \tag{4-30}$$

Therefore, we have to enhance the exponent  $\frac{2}{3}$  on the right-hand side to 1; as a sacrifice, the Hölder exponent on the left-hand side will decrease.

Set  $\tau_y g(w) := g(w \circ (0, y, 0))$  for  $y \in \mathbb{R}^d$  and  $w \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d$ . A direct computation shows that  $q$  satisfies

$$(\partial_t + v \cdot \nabla_x)q = \mathcal{R}[g]\mathcal{U}[q] + \sum_{\substack{|i| \leq n \\ i \leq k_x}} \delta_y \hat{D}_i \mathcal{R}[g]\mathcal{U}[\tau_y D_i g] + \sum_{\substack{|i| \leq n-1 \\ i \leq k_x}} \hat{D}_i \mathcal{R}[g]\mathcal{U}[\delta_y D_i g], \tag{4-31}$$

where the multi-indices  $i \leq k_x$  mean each component of  $i$  is less than or equal to the corresponding component of  $k$ , and  $\hat{D}_i$  denotes the differential operator satisfying  $\partial_t^{k_t} D_x^{k_x} = \hat{D}_i \circ D_i$ .

In view of (4-29), (4-30), and the induction hypothesis, each term in the summations on the right-hand side of (4-31) is bounded in  $C_t^{\alpha_n}$  by  $C_n \|(0, y, 0)\|^2 \mu^{v'}(v_0)$  for any  $v' \in (0, v)$ . Then, by the interior Schauder estimate (Proposition 3.1),

$$\|q\|_{C_t^{2+\alpha_n}} \lesssim_{T,v',n} \|(0, y, 0)\|^2 \mu^{v'}(v_0). \tag{4-32}$$

Combining Lemma C.1 with (4-30) and (4-32), we obtain (4-28).

Step 2. For the case  $k_v = 0$  in (4-27), we proceed with a bidimensional induction on  $(m, n) = (k_t, |k_x|)$  such that, for any  $v \in (0, \frac{1}{2})$ ,

$$\|\partial_t^{k_t} D_x^{k_x} g\|_{C_t^{2+\alpha_{m,n}}} \lesssim_{T,v,m,n} \mu^v(v_0). \tag{4-33}$$

Based on the previous step ( $m = 0$ ), we have to show that (4-33) holds for  $k_t = m \geq 1$  and  $|k_x| = n$  under the induction hypothesis that (4-33) holds for any  $k_t \leq m - 1$  and  $|k_x| \leq n + 1$ .

With  $k_t = m > 0$  and  $|k_x| = n$ , set  $q := \partial_t^{k_t} D_x^{k_x} g$ . Then,

$$(\partial_t + v \cdot \nabla_x)q = \mathcal{R}[g]\mathcal{U}[q] + \sum_{\substack{i \leq (k_t, k_x, 0) \\ i \neq (k_t, k_x, 0)}} \hat{D}_i \mathcal{R}[g]\mathcal{U}[D_i g], \tag{4-34}$$

where we use the notation  $\hat{D}_i$  for the differential operator satisfying  $\partial_t^{k_t} D_x^{k_x} = \hat{D}_i \circ D_i$ .

By the induction hypothesis, each term in the remainder (the summation on the right-hand side of (4-34)) with  $i \neq (0, 0, 0)$  can be controlled in  $C_l^{\alpha_{m,n}}$ . It now suffices to deal with the exceptional term  $\partial_t^{k_t} D_x^{k_x} \mathcal{R}[g]$  so that the whole remainder can be controlled in  $C_l^{\alpha_{m,n}}$ ; then (4-33) follows from the interior Schauder estimate (Proposition 3.1). To this end, using Lemma 2.4 and the induction hypothesis with the pair  $(m - 1, n)$  yields

$$\|(\partial_t + v \cdot \nabla_x)\partial_t^{m-1} D_x^{k_x} g\|_{C_l^{\alpha_{m,n}}} \lesssim_{\underline{T}, v, m, n} \mu^v(v_0). \tag{4-35}$$

Due to the induction hypothesis with the pair  $(m - 1, n + 1)$ , for any  $v' \in (0, v)$ ,

$$\mu^{-v'}(v_0) \|(v \cdot \nabla_x)\partial_t^{m-1} D_x^{k_x} g\|_{C_l^{2+\alpha_{m,n}}} \lesssim_{v, v'} \mu^{-v}(v_0) \|\partial_t^{m-1} \nabla_x D_x^{k_x} g\|_{C_l^{2+\alpha_{m,n}}} \lesssim_{\underline{T}, v, m, n} 1. \tag{4-36}$$

Then, (4-35) and (4-36) produce the bound on  $\mu^{-v'}(v_0) \|q\|_{C_l^{\alpha_{m,n}}}$ .

Step 3. Similarly, to show (4-27) for any differential operator  $\partial_t^{k_t} D_x^{k_x} D_v^{k_v}$ , we proceed with a bidimensional induction on  $(m, n) = (k_t + |k_x|, k_v)$  such that, for any  $v \in (0, \frac{1}{2})$ ,

$$\|\partial_t^{k_t} D_x^{k_x} D_v^{k_v} g\|_{C_l^{2+\alpha_{m,n}}} \lesssim_{\underline{T}, v, m, n} \mu^v(v_0). \tag{4-37}$$

The case  $n = 0$  is treated in the previous step. By Lemma 2.4 and the induction hypothesis (4-37) with  $k_t + |k_x| = m$  and  $|k_v| = n - 1, n \geq 1$ , we have

$$\|\partial_v \partial_t^{k_t} \partial_x^{k_x} \partial_v^{k_v} g\|_{C_l^{\alpha_{m,n}}} \lesssim \|\partial_t^{k_t} \partial_x^{k_x} \partial_v^{k_v} g\|_{C_l^{1+\alpha_{m,n}}} \lesssim_{\underline{T}, v, m, n} \mu^v(v_0).$$

Computing the equation satisfied by  $\partial_v \partial_t^{k_t} \partial_x^{k_x} \partial_v^{n-1} g$  and proceeding as in the previous step, we conclude the proof. □

### 5. Diffusion asymptotics

This section is devoted to the study of the global-in-time quantitative diffusion asymptotics which consists of the (uniform-in- $\epsilon$ ) convergence towards the equilibrium over long times and of the finite-time asymptotics, including the results of Theorem 1.1(ii) and Theorem 1.4.

We first introduce the required notation. For any scalar or vector-valued function  $\Psi \in L^1(\mathbb{R}^d, d\mu)$ , we denote its velocity mean by

$$\langle \Psi \rangle := \int_{\mathbb{R}^d} \Psi(v) d\mu.$$

For any pair of functions (scalars, vectors, or  $d \times d$  matrices)  $\Psi_1, \Psi_2 \in L^2(\mathbb{T}^d \times \mathbb{R}^d, dm)$ , we denote their  $L^2$  inner product with respect to the measure  $dm$  by

$$(\Psi_1, \Psi_2) := \int_{\mathbb{T}^d \times \mathbb{R}^d} \Psi_1(x, v) \Psi_2(x, v) dm,$$

where the multiplication between the pair in the integrand is replaced by the scalar contraction product if  $\Psi_1$  and  $\Psi_2$  are a pair of vectors or matrices.

Recalling our notation for the Ornstein–Uhlenbeck operator  $\mathcal{L}_{OU} = (\nabla_v - v) \cdot \nabla_v$ , we apply the substitutions  $f_\epsilon = \mu h_\epsilon$  and  $f_{\epsilon, \text{in}} = \mu h_{\epsilon, \text{in}}$  in (1-2) and obtain

$$\begin{cases} (\epsilon \partial_t + v \cdot \nabla_x) h_\epsilon(t, x, v) = \frac{1}{\epsilon} \langle h_\epsilon \rangle^\beta(t, x) \mathcal{L}_{OU} h_\epsilon(t, x, v), \\ h_\epsilon(0, x, v) = h_{\epsilon, \text{in}}(x, v). \end{cases} \tag{5-1}$$

In this setting, by applying integration by parts, for any  $h_1, h_2 \in C_c^\infty(\mathbb{T}^d \times \mathbb{R}^d)$ , we get

$$(h_1, \mathcal{L}_{OU} h_2) = -(\nabla_v h_1, \nabla_v h_2).$$

We will use this identity repeatedly in the computation below. Then, the operator  $\mathcal{L}_{OU}$  is self-adjoint with respect to the inner product  $(\cdot, \cdot)$ , and the bracket  $\langle \cdot \rangle$  is a projection on the null space of  $\mathcal{L}_{OU}$ . Moreover, as the total mass is conserved, we define

$$M_0 := \int_{\mathbb{T}^d \times \mathbb{R}^d} h_\epsilon dm = \int_{\mathbb{T}^d} \langle h_\epsilon \rangle dx. \tag{5-2}$$

Proceeding with the macro-micro (fluid-kinetic) decomposition, we define the orthogonal complement of the projection  $\langle \cdot \rangle$  of  $h_\epsilon$  as

$$h_\epsilon^\perp(t, x, v) := h_\epsilon(t, x, v) - \langle h_\epsilon \rangle(t, x).$$

In this framework, the local mass  $\langle h_\epsilon \rangle$  is the macroscopic (fluid) part and the complement  $h_\epsilon^\perp$  is the microscopic (kinetic) part. In addition, taking the bracket  $\langle \cdot \rangle$  after multiplying the equation in (5-1) with 1 and  $v$  leads to the macroscopic equations

$$\epsilon \partial_t \langle h_\epsilon \rangle + \nabla_x \cdot \langle v h_\epsilon \rangle = 0, \tag{5-3}$$

$$\epsilon \partial_t \langle v h_\epsilon \rangle + \nabla_x \cdot \langle v^{\otimes 2} h_\epsilon \rangle = -\frac{1}{\epsilon} \langle h_\epsilon \rangle^\beta \langle v h_\epsilon \rangle, \tag{5-4}$$

where  $\langle v h_\epsilon \rangle$  and  $\langle v^{\otimes 2} h_\epsilon \rangle$  represent the local momentum and the stress tensor, respectively.

**5A. Long time behavior.** Our aim is to establish the (uniform-in- $\epsilon$ ) exponential decay towards the equilibrium  $M_0$  for (5-1). In particular, when  $\epsilon = 1$ , it sets up the exponential convergence in each order derivative based on the smoothness a priori estimates given in Section 4C.

We note that the classical coercive method is not applicable in our case to obtain the convergence to equilibrium due to the degeneracy of the ellipticity of the spatially inhomogeneous equation. Indeed, the Poincaré inequality only produces a spectral gap on the orthogonal complement of the projection  $\langle \cdot \rangle$ ; see (5-7). As mentioned in Section 1B, there are several ways to achieve the long-time asymptotics. We mainly follow the argument in [Esposito et al. 2013] (see also [Kim et al. 2020]) in a simpler scenario. It also allows us to see some similarity among [Dolbeault et al. 2015; Esposito et al. 2013; Hérau 2018].

**Proposition 5.1.** *Let the function  $\lambda_t : \mathbb{R}_+ \rightarrow [0, \Lambda]$  satisfy  $\lambda'_t \leq 0$  on  $\mathbb{R}_+$ . If  $h_\epsilon$  is a solution to (5-1) in  $\mathbb{R}_+ \times \mathbb{T}^d \times \mathbb{R}^d$  associated with the initial data  $0 \leq h_{\epsilon, \text{in}} \leq \Lambda$  and satisfying*

$$\langle h_\epsilon \rangle^\beta(t, x) \geq \lambda_t \quad \text{in } \mathbb{R}_+ \times \mathbb{T}^d \quad \text{and} \quad \int_{\mathbb{R}_+} (\lambda_t + \lambda'_t) dt = \infty, \tag{5-5}$$

*then the solution  $h_\epsilon$  converges to the state  $M_0$  in  $L^2(dm)$  as  $t \rightarrow \infty$ . More precisely, there exists some universal constant  $c > 0$  such that, for any  $t > 0$ , we have*

$$\|h_\epsilon(t, \cdot, \cdot) - M_0\|_{L^2(dm)}^2 \lesssim \|h_{\epsilon, \text{in}} - M_0\|_{L^2(dm)}^2 \exp\left(-c \int_0^t (\lambda_s + \lambda'_s) ds\right). \tag{5-6}$$

*Proof.* Since the velocity mean of the microscopic part vanishes,  $\langle h_\epsilon^\perp \rangle = 0$ , using (5-1) and the Poincaré inequality yields

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|h_\epsilon - M_0\|_{L^2(dm)}^2 &= \frac{1}{\epsilon^2} (\langle h_\epsilon \rangle^\beta \mathcal{L}_{\text{OU}} h_\epsilon, h_\epsilon - M_0) = -\frac{1}{\epsilon^2} (\langle h_\epsilon \rangle^\beta \nabla_v h_\epsilon^\perp, \nabla_v h_\epsilon^\perp) \\ &\leq -\frac{\lambda_t}{\epsilon^2} \|\nabla_v h_\epsilon^\perp\|_{L^2(dm)}^2 \lesssim -\frac{\lambda_t}{\epsilon^2} \|h_\epsilon^\perp\|_{L^2(dm)}^2. \end{aligned} \tag{5-7}$$

Now we have to recover a new entropy that would give some bound on the projection  $\langle h_\epsilon \rangle - M_0$ .

For every test function  $v \cdot \Psi(t, x)\mu$ , with a vector-valued function  $\Psi \in H^1_{t,x}(\mathbb{R}^+ \times \mathbb{T}^d, \mathbb{R}^d)$ , we write the weak formulation of (5-1) as

$$\frac{d}{dt} (v \cdot \Psi, h_\epsilon) = \frac{1}{\epsilon} (v^{\otimes 2} : \nabla_x \Psi, h_\epsilon) + (v \cdot \partial_t \Psi, h_\epsilon) + \frac{1}{\epsilon^2} (\langle h_\epsilon \rangle^\beta \mathcal{L}_{\text{OU}} v \cdot \Psi, h_\epsilon).$$

Taking the macro-micro decomposition into account, from the above expression we obtain

$$\begin{aligned} \frac{d}{dt} (v \cdot \Psi, h_\epsilon^\perp) &= \frac{1}{\epsilon} (|v_1|^2 \text{tr}(\nabla_x \Psi), \langle h_\epsilon \rangle - M_0) + \frac{1}{\epsilon} (v^{\otimes 2} : \nabla_x \Psi, h_\epsilon^\perp) \\ &\quad + (v \cdot \partial_t \Psi, h_\epsilon^\perp) - \frac{1}{\epsilon^2} (\langle h_\epsilon \rangle^\beta v \cdot \Psi, h_\epsilon^\perp). \end{aligned} \tag{5-8}$$

Let us now introduce an auxiliary function  $u(t, x)$ : for any fixed  $t \in \mathbb{R}_+$ , defined  $u(t, x)$  as the solution of the following elliptic equation under the compatibility condition (5-2):

$$-\Delta_x u = \langle h_\epsilon \rangle - M_0 \quad \text{in } \mathbb{T}^d, \tag{5-9}$$

whose elliptic estimate states

$$\|\nabla_x u\|_{L^2_x} + \|\nabla_x^2 u\|_{L^2_x} \lesssim \|\langle h_\epsilon \rangle - M_0\|_{L^2_x}. \tag{5-10}$$

In addition, observing that  $\langle v h_\epsilon \rangle = \langle v h_\epsilon^\perp \rangle$ , from (5-3) we get

$$\epsilon \partial_t \langle h_\epsilon \rangle + \nabla_x \cdot \langle v h_\epsilon^\perp \rangle = 0.$$

Combining this macroscopic relation with (5-9), we have

$$\int_{\mathbb{T}^d} |\nabla_x(\partial_t u)|^2 = \int_{\mathbb{T}^d} \partial_t u \partial_t \langle h_\epsilon \rangle = -\frac{1}{\epsilon} \int_{\mathbb{T}^d} \partial_t u \nabla_x \cdot \langle v h_\epsilon^\perp \rangle = \frac{1}{\epsilon} \int_{\mathbb{T}^d} \nabla_x(\partial_t u) \cdot \langle v h_\epsilon^\perp \rangle.$$

It then follows from Hölder's inequality that

$$\|\nabla_x(\partial_t u)\|_{L_x^2} \leq \frac{1}{\epsilon} \|\langle v h_\epsilon^\perp \rangle\|_{L_x^2} \lesssim \frac{1}{\epsilon} \|h_\epsilon^\perp\|_{L^2(dm)}. \quad (5-11)$$

Choosing  $\Psi = \nabla_x u$  in (5-8) yields

$$-\frac{1}{\epsilon} (|v_1|^2 \Delta_x u, \langle h_\epsilon \rangle - M_0) + \frac{d}{dt} (v \cdot \nabla_x u, h_\epsilon^\perp) \lesssim \left( \frac{1}{\epsilon} \|\nabla_x^2 u\|_{L_x^2} + \|\nabla_x(\partial_t u)\|_{L_x^2} + \frac{1}{\epsilon^2} \|\nabla_x u\|_{L_x^2} \right) \|h_\epsilon^\perp\|_{L^2(dm)}.$$

Applying (5-9)–(5-11), we have

$$\frac{1}{\epsilon} \|\langle h_\epsilon \rangle - M_0\|_{L_x^2}^2 + \frac{d}{dt} (v \cdot \nabla_x u, h_\epsilon^\perp) \lesssim \frac{1}{\epsilon^2} \|\langle h_\epsilon \rangle - M_0\|_{L_x^2} \|h_\epsilon^\perp\|_{L^2(dm)} + \frac{1}{\epsilon} \|h_\epsilon^\perp\|_{L^2(dm)}^2.$$

By the Cauchy–Schwarz inequality, we arrive at

$$\|\langle h_\epsilon \rangle - M_0\|_{L_x^2}^2 + \epsilon \frac{d}{dt} (v \cdot \nabla_x u, h_\epsilon^\perp) \lesssim \frac{1}{\epsilon^2} \|h_\epsilon^\perp\|_{L^2(dm)}^2. \quad (5-12)$$

Then, (5-12) combined with (5-7) implies

$$\begin{aligned} \frac{d}{dt} \mathcal{E}_\epsilon(t) &\lesssim -\frac{1-\delta}{\epsilon^2} \|h_\epsilon^\perp\|_{L^2(dm)}^2 - \delta \lambda_t \|\langle h_\epsilon \rangle - M_0\|_{L_x^2}^2 + \delta \epsilon \lambda_t' (v \cdot \nabla_x u, h_\epsilon^\perp) \\ &\leq -\delta \lambda_t \|h_\epsilon - M_0\|_{L^2(dm)}^2 - \delta \lambda_t' |(v \cdot \nabla_x u, h_\epsilon^\perp)|, \end{aligned}$$

where the constant  $\delta \in (0, \frac{1}{2})$  will be determined and the modified entropy  $\mathcal{E}_\epsilon$  is defined by

$$\mathcal{E}_\epsilon(t) := \|h_\epsilon - M_0\|_{L^2(dm)}^2 + \delta \epsilon \lambda_t (v \cdot \nabla_x u, h_\epsilon^\perp).$$

We note that (5-10) also implies

$$|(v \cdot \nabla_x u, h_\epsilon^\perp)| \lesssim \|\langle h_\epsilon \rangle - M_0\|_{L_x^2} \|h_\epsilon^\perp\|_{L^2(dm)} \leq \|h_\epsilon - M_0\|_{L^2(dm)}^2. \quad (5-13)$$

It means that the modified entropy  $\mathcal{E}_\epsilon$  is equivalent (independent of  $\epsilon$ ) to the square of the  $L^2(dm)$ -distance between  $h_\epsilon$  and  $M_0$ , when the constant  $\delta > 0$  is sufficiently small.

Hence we have

$$\frac{d}{dt} \mathcal{E}_\epsilon(t) \lesssim -(\lambda_t + \lambda_t') \mathcal{E}_\epsilon(t).$$

The conclusion (5-6) then follows from Grönwall's inequality and the equivalence between  $\mathcal{E}_\epsilon(t)$  and  $\|h_\epsilon(t, \cdot, \cdot) - M_0\|_{L^2(dm)}^2$ .  $\square$

We pointed out that the elliptic estimate (5-10) for the Poisson equation (5-9) used in the above proof resulting from a Poincaré-type inequality essentially relies on the compactness of the spatial domain. It was shown in [Bouin et al. 2020] that the related elliptic estimate can be recovered by applying the Nash inequality [1958] when the spatial domain is the whole space  $\mathbb{R}^d$ , whose argument is under an abstract setting. Inspired by the proof of Proposition 5.1 above, we are also able to make the construction of [Bouin et al. 2020] precise to see that the argument still works for the nonlinear equation (5-1). We remark that the following algebraic decay rate is optimal in the sense that it is the same as in the linear case; see Appendix A of [Bouin et al. 2020].

**Proposition 5.2.** *Assume the initial data  $h_{\epsilon, \text{in}}$  is valued in  $[\lambda, \Lambda]$  and satisfies  $h_{\epsilon, \text{in}} - M_1 \in L^1(\mathbb{R}^{2d}, dm)$  for some universal constant  $M_1 > 0$ . Let the function  $h_\epsilon$  valued in  $[\lambda, \Lambda]$  be a solution to (5-1) in  $\mathbb{R}_+ \times \mathbb{R}^{2d}$  associated with  $h_\epsilon|_{t=0} = h_{\epsilon, \text{in}}$ . Then, for any  $t > 0$ ,*

$$\|h_\epsilon - M_1\|_{L^2(\mathbb{R}^{2d}, dm)} \lesssim (1 + \|h_{\epsilon, \text{in}} - M_1\|_{L^1(\mathbb{R}^{2d}, dm)})t^{-d/4}.$$

*Proof.* By the same derivation of (5-7) and (5-8) as in the proof of Proposition 5.1, we have the microscopic coercivity

$$\frac{d}{dt} \|h_\epsilon - M_1\|_{L^2(\mathbb{R}^{2d}, dm)}^2 \lesssim -\frac{1}{\epsilon^2} \|h_\epsilon^\perp\|_{L^2(\mathbb{R}^{2d}, dm)}^2, \tag{5-14}$$

and the identity from the macro-micro decomposition

$$\begin{aligned} & -(|v_1|^2 \Delta_x w, \langle h_\epsilon \rangle - M_0)_W + \epsilon \frac{d}{dt} (v \cdot \nabla_x w, h_\epsilon^\perp)_W \\ & = (v^{\otimes 2} : \nabla_x^2 w, h_\epsilon^\perp)_W + \epsilon (v \cdot \partial_t \nabla_x w, h_\epsilon^\perp)_W - \frac{1}{\epsilon} (\langle h_\epsilon \rangle^\beta v \cdot \nabla_x w, h_\epsilon^\perp)_W, \end{aligned} \tag{5-15}$$

where  $(\cdot, \cdot)_W$  denotes the  $L^2(\mathbb{R}^{2d}, dm)$  inner product, and the function  $w(t, x) \in L_t^\infty([0, T]; L_x^1 \cap L_x^2(\mathbb{R}^d))$  is chosen to be the solution of the following elliptic equation associated with the constant  $\Theta := \langle |v_1|^2 \rangle$  and the macroscopic source  $\langle h_\epsilon \rangle - M_1$ :

$$w - \Theta \Delta_x w = \langle h_\epsilon \rangle - M_1 \quad \text{in } \mathbb{R}^d. \tag{5-16}$$

The elliptic estimate is derived by integrating (5-16) against  $-\Theta \Delta_x w$ , so that

$$\begin{aligned} \Theta \|\nabla_x w\|_{L_x^2(\mathbb{R}^d)}^2 + \Theta^2 \|\nabla_x^2 w\|_{L_x^2(\mathbb{R}^d)}^2 & = (-\Theta \Delta_x w, \langle h_\epsilon \rangle - M_1)_W \\ & = (\langle h_\epsilon \rangle - M_1 - w, \langle h_\epsilon \rangle - M_1)_W =: \mathcal{A}. \end{aligned} \tag{5-17}$$

It also follows from the same derivation as (5-11) that

$$\|\nabla_x(\partial_t w)\|_{L_x^2(\mathbb{R}^d)} \lesssim \frac{1}{\epsilon} \|h_\epsilon^\perp\|_{L^2(\mathbb{R}^{2d}, dm)}.$$

Combining the above two estimates with (5-15), we obtain

$$\mathcal{A} + \epsilon \frac{d}{dt} (v \cdot \nabla_x w, h_\epsilon^\perp)_W \lesssim \frac{1}{\epsilon} \mathcal{A}^{1/2} \|h_\epsilon^\perp\|_{L^2(\mathbb{R}^{2d}, dm)} + \|h_\epsilon^\perp\|_{L^2(\mathbb{R}^{2d}, dm)}^2,$$

which implies from the Cauchy–Schwarz inequality that

$$\mathcal{A} + \epsilon \frac{d}{dt} (v \cdot \nabla_x w, h_\epsilon^\perp)_W \lesssim \frac{1}{\epsilon^2} \|h_\epsilon^\perp\|_{L^2(\mathbb{R}^{2d}, dm)}^2.$$

Denoting the modified entropy by

$$\mathcal{E}_\epsilon(t) := \|h_\epsilon - M_1\|_{L^2(\mathbb{R}^{2d}, dm)}^2 + \delta \epsilon (v \cdot \nabla_x w, h_\epsilon^\perp)_W$$

with a sufficiently small constant  $\delta > 0$  and using (5-14), we conclude that

$$\frac{d}{dt} \mathcal{E}_\epsilon(t) \lesssim -\mathcal{A} - \|h_\epsilon^\perp\|_{L^2(\mathbb{R}^{2d}, dm)}^2. \tag{5-18}$$

Recalling the similar estimate (5-13), we see that  $\mathcal{E}_\epsilon$  is equivalent to the square of the  $L^2(\mathbb{R}^{2d}, dm)$ -distance between  $h_\epsilon$  and  $M_1$ . It thus suffices to recover  $\langle h_\epsilon \rangle - M_1$  by means of  $\mathcal{A}$ . By (5-16) and the convexity of  $|\cdot|$ , we know that  $|w|$  is a subsolution in the sense that

$$|w| - \Theta \Delta_x |w| \leq |\langle h_\epsilon \rangle - M_1| \quad \text{in } \mathbb{R}^d,$$

and hence  $\|w\|_{L^1_x(\mathbb{R}^d)} \leq \|\langle h_\epsilon \rangle - M_1\|_{L^1_x(\mathbb{R}^d)}$ . With the aid of Corollary 4.10,

$$\|w\|_{L^1_x(\mathbb{R}^d)} \leq \|h - M_1\|_{L^1(\mathbb{R}^{2d}, dm)} \leq \|h_{\epsilon, \text{in}} - M_1\|_{L^1(\mathbb{R}^{2d}, dm)}.$$

Applying (5-17) and the Nash inequality  $\|w\|_{L^2_x(\mathbb{R}^d)}^{d+2} \lesssim \|w\|_{L^1_x(\mathbb{R}^d)}^2 \|\nabla_x w\|_{L^2_x(\mathbb{R}^d)}^d$ , we then acquire

$$\begin{aligned} \|\langle h_\epsilon \rangle - M_1\|_{L^2_x(\mathbb{R}^d)}^2 &\lesssim \mathcal{A} + \|w\|_{L^2_x(\mathbb{R}^d)}^2 \lesssim \mathcal{A} + \|h_{\epsilon, \text{in}} - M_1\|_{L^1(\mathbb{R}^{2d}, dm)}^{4/(d+2)} \|\nabla_x w\|_{L^2_x(\mathbb{R}^d)}^{2d/(d+2)} \\ &\lesssim (\mathcal{A}^{2/(d+2)} + \|h_{\epsilon, \text{in}} - M_1\|_{L^1(\mathbb{R}^{2d}, dm)}^{4/(d+2)}) \mathcal{A}^{d/(d+2)}. \end{aligned}$$

Since  $\|h_\epsilon - M_1\|_{L^2_x(\mathbb{R}^d)}^2 \leq (\Lambda + M_1) \|h_{\epsilon, \text{in}} - M_1\|_{L^1(\mathbb{R}^{2d}, dm)}$ , in both cases

$$\mathcal{A} + \|h_\epsilon^\perp\|_{L^2(\mathbb{R}^{2d}, dm)}^2 \lesssim \|h_\epsilon - M_1\|_{L^2(\mathbb{R}^{2d}, dm)}^2,$$

we conclude that

$$\|h_\epsilon - M_1\|_{L^2(\mathbb{R}^{2d}, dm)}^2 \lesssim (1 + \|h_{\epsilon, \text{in}} - M_1\|_{L^1(\mathbb{R}^{2d}, dm)}^{4/(d+2)}) (\mathcal{A} + \|h_\epsilon^\perp\|_{L^2(\mathbb{R}^{2d}, dm)}^2)^{d/(d+2)}.$$

Combining this with (5-18) and the equivalence between  $\mathcal{E}_\epsilon$  and  $\|h_\epsilon - M_1\|_{L^2(\mathbb{R}^{2d}, dm)}^2$ , we have

$$\frac{d}{dt} \mathcal{E}_\epsilon(t) \lesssim -(1 + \|h_{\epsilon, \text{in}} - M_1\|_{L^1(\mathbb{R}^{2d}, dm)}^{4/d})^{-1} \mathcal{E}_\epsilon(t)^{1+2/d},$$

Since  $\mathcal{E}_\epsilon(0) \lesssim \|h_{\epsilon, \text{in}} - M_1\|_{L^1(\mathbb{R}^{2d}, dm)}$ , we arrive at

$$\mathcal{E}_\epsilon(t) \lesssim [\mathcal{E}_\epsilon(0)^{-2/d} + (1 + \|h_{\epsilon, \text{in}} - M_1\|_{L^1(\mathbb{R}^{2d}, dm)}^{4/d})^{-1} t]^{-d/2} \lesssim (1 + \|h_{\epsilon, \text{in}} - M_1\|_{L^1(\mathbb{R}^{2d}, dm)}^2) t^{-d/2}. \quad \square$$

As far as the case  $\epsilon = 1$  is concerned, we conclude the result of convergence to equilibrium.

*Proof of Theorem 1.1(ii).* Consider  $g := \mu^{1/2}h$ . In view of Proposition 4.8 and Corollary 4.10 with the assumption on initial data, we know that  $\lambda \leq \mu^{-1/2}g \leq \Lambda$  in  $\mathbb{R}_+ \times \Omega_x \times \mathbb{R}^d$  for  $\Omega_x = \mathbb{T}^d$  or  $\mathbb{R}^d$ . By applying Proposition 5.1 to  $h = \mu^{-1/2}g$  with  $\lambda_t = \lambda$  and  $\Omega_x = \mathbb{T}^d$ , we have an universal constant  $c > 0$  such that

$$\|g(t) - M_0 \mu^{1/2}\|_{L^2(\mathbb{T}^d \times \mathbb{R}^d)} \lesssim e^{-2ct}.$$

Combining this with the Sobolev embedding and the interpolation, we derive the following for any  $k \in \mathbb{N}$  with  $k \geq d$ :

$$\begin{aligned} \|g(t) - M_0 \mu^{1/2}\|_{C^k(\mathbb{T}^d \times \mathbb{R}^d)} &\lesssim_k \|g(t) - M_0 \mu^{1/2}\|_{H^{2k}(\mathbb{T}^d \times \mathbb{R}^d)} \\ &\lesssim_k \|g(t) - M_0 \mu^{1/2}\|_{H^{4k}(\mathbb{T}^d \times \mathbb{R}^d)}^{1/2} \|g(t) - M_0 \mu^{1/2}\|_{L^2(\mathbb{T}^d \times \mathbb{R}^d)}^{1/2}. \end{aligned}$$

Since the  $H^{4k}$ -norm on the right-hand side is bounded due to the global regularity estimate given by Proposition 4.12, we obtain the exponential convergence to equilibrium in each order derivative.

The asserted result in the case  $\Omega_x = \mathbb{R}^d$  is a direct consequence of [Proposition 5.2](#). As a side remark, one is also able to upgrade the long-time convergence to higher-order derivatives of solutions by means of the global regularity estimate and interpolation as above; it yet gives the algebraic decay rate that is not optimal.  $\square$

**5B. Finite-time asymptotics.** The study of macroscopic dynamics for the nonlinear kinetic model (5-1) in this subsection relies on the regularity of the target equation (1-3). On account of this, let us begin with mentioning some standard results for (1-3) without proof. If the initial data satisfies  $\lambda \leq \rho_{\text{in}} \leq \Lambda$ , then such bounds are preserved along times,  $\lambda \leq \rho \leq \Lambda$ , in the same spirit as [Lemma 4.1](#). Combining the parabolic De Giorgi–Nash–Moser theory with Schauder theory, we know that the solution  $\rho$  is smooth for any positive time. We state the a priori estimate precisely as follows, where its behavior near the initial time is taken into account in view of the standard scaling technique.

**Lemma 5.3.** *Let  $\rho_{\text{in}} \in C^{\alpha_0}(\mathbb{T}^d)$  be valued in  $[\lambda, \Lambda]$  with  $\alpha_0 \in (0, 1)$ , and let  $\rho$  be the solution to (1-3) in  $\mathbb{R}_+ \times \mathbb{T}^d$ . Then there is some universal constant  $\alpha \in (0, 1)$  such that*

$$\|\rho\|_{C^\alpha(\mathbb{R}_+ \times \mathbb{T}^d)} \lesssim 1 + \|\rho_{\text{in}}\|_{C^{\alpha_0}(\mathbb{T}^d)}. \tag{5-19}$$

Moreover, there exists some constant  $C_\rho > 0$  depending only on universal constants and  $\|\rho_{\text{in}}\|_{C^{\alpha_0}(\mathbb{T}^d)}$  such that, for any  $t \in (0, 1]$  and  $x \in \mathbb{T}^d$ , we have

$$t^{(1-\alpha)/2} |\nabla_x \rho(t, x)| + t^{(2-\alpha)/2} |\partial_t \rho(t, x)| + t^{(2-\alpha)/2} |\nabla_x^2 \rho(t, x)| + t^{(3-\alpha)/2} |\partial_t \nabla_x \rho(t, x)| \leq C_\rho, \tag{5-20}$$

and, for any  $t \geq 1$ , we have

$$\|\nabla_x \rho(t, \cdot)\|_{L^\infty(\mathbb{T}^d)} + \|\partial_t \rho(t, \cdot)\|_{L^\infty(\mathbb{T}^d)} + \|\nabla_x^2 \rho(t, \cdot)\|_{L^\infty(\mathbb{T}^d)} + \|\partial_t \nabla_x \rho(t, \cdot)\|_{L^\infty(\mathbb{T}^d)} \lesssim 1. \tag{5-21}$$

We measure the distance between solutions to the scaled nonlinear kinetic model (1-2) and solutions to the fast diffusion equation (1-3) by the relative phi-entropy functional  $\mathcal{H}_\beta$  (see [Definition 1.3](#)). The following lemma shows the effectiveness of the relative phi-entropy for measuring  $L^2$ -distance by virtue of the uniform convexity of  $\varphi_\beta$ . It can be seen as a simple version of the Csiszár–Kullback inequality on the relative entropy. We give its statement below with a proof taken from [\[Dolbeault and Li 2018, Proposition 2.1\]](#) for the sake of completeness.

**Lemma 5.4.** *Let  $h_1$  and  $h_2$  be two functions valued in  $[0, \Lambda]$ . Then we have*

$$\mathcal{H}_\beta(h_1 | h_2) \geq \left(1 - \frac{1}{2}\beta\right) \Lambda^{-\beta} \|h_1 - h_2\|_{L^2(dm)}^2. \tag{5-22}$$

If we additionally assume the lower bound  $h_1, h_2 \geq \lambda$ , then

$$\mathcal{H}_\beta(h_1 | h_2) \leq \left(1 - \frac{1}{2}\beta\right) \lambda^{-\beta} \|h_1 - h_2\|_{L^2(dm)}^2.$$

*Proof.* Since  $\varphi_\beta(1) = \varphi'_\beta(1) = 0$  and  $\beta \in [0, 1]$ , for any  $z \in \mathbb{R}_+$ , there exists  $\xi_z \in \mathbb{R}_+$  lying between 1 and  $z$  such that

$$\varphi_\beta(z) = \frac{1}{2} \varphi''_\beta(\xi_z) (z - 1)^2 = \frac{1}{2} (2 - \beta) \xi_z^{-\beta} (z - 1)^2.$$



Since  $\min\{z, 1\} \leq \xi_z \leq \max\{z, 1\}$ , we have

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} \max\{h_1^{-\beta}, h_2^{-\beta}\} |h_1 - h_2|^2 \, dm \leq \frac{2}{2-\beta} \mathcal{H}_\beta(h_1 | h_2) \leq \int_{\mathbb{T}^d \times \mathbb{R}^d} \min\{h_1^{-\beta}, h_2^{-\beta}\} |h_1 - h_2|^2 \, dm,$$

which implies the desired results by using the boundedness of  $h_1$  and  $h_2$ . □

Let us consider the finite-time diffusion asymptotics.

**Proposition 5.5.** *Let  $\rho_{\text{in}} \in C^{\alpha_0}(\mathbb{T}^d)$  be valued in  $[\lambda, \Lambda]$  with  $\alpha_0 \in (0, 1)$ , and let the sequence of functions  $\{h_{\epsilon, \text{in}}\}_{\epsilon \in (0, 1)} \subset C^{\alpha_0}(\mathbb{T}^d \times \mathbb{R}^d)$  satisfy*

$$\langle h_{\epsilon, \text{in}} \rangle \geq \lambda \quad \text{in } \mathbb{T}^d \quad \text{and} \quad 0 \leq h_{\epsilon, \text{in}} \leq \Lambda \quad \text{in } \mathbb{T}^d \times \mathbb{R}^d.$$

Let  $h_\epsilon$  be the solutions of (5-1) associated with this initial data. Then there exist some universal constants  $\alpha \in (0, 1)$  and  $C > 0$ , and some constant  $C_\rho > 0$  depending only on universal constants and  $\|\rho_{\text{in}}\|_{C^{\alpha_0}(\mathbb{T}^d)}$  such that, for any  $\epsilon \in (0, 1)$  and for any  $t \in [\underline{T}, 1]$  with  $\underline{T} \in (0, 1)$ , the following estimate holds:

$$\mathcal{H}_\beta(h_\epsilon | \rho)(t) \leq C_\rho \mathcal{H}_\beta(h_\epsilon | \rho)(\underline{T}) + C_\rho \epsilon (t^{(\alpha-1)/2} + \epsilon t^{(\alpha-2)/2}), \tag{5-23}$$

where  $\rho(t, x)$  is the solution to (1-3) associated with the initial data  $\rho_{\text{in}}$ , and, for any  $t \geq 1$ , we have

$$\mathcal{H}_\beta(h_\epsilon | \rho)(t) \leq [\mathcal{H}_\beta(h_\epsilon | \rho)(t)(1) + C\epsilon(1 + t^{1/2})]e^{Ct}. \tag{5-24}$$

*Proof.* For  $\beta \in [0, 1)$ , the phi-entropy of  $h_\epsilon$  relative to  $\rho$  reads

$$\mathcal{H}_\beta(h_\epsilon | \rho) = \mathcal{H}_\beta(h_\epsilon | 1) - \mathcal{H}_\beta(\rho | 1) - \frac{2-\beta}{1-\beta} (\langle h_\epsilon \rangle - \rho, \rho^{1-\beta} - 1).$$

As far as the entropy  $\mathcal{H}_\beta(h_\epsilon | 1)$  is concerned, the entropy dissipation is derived by (5-1), integration by parts, and using Hölder's inequality  $\langle \nabla_v h_\epsilon \rangle^2 \leq \langle h_\epsilon \rangle^\beta \langle h_\epsilon^{-\beta} |\nabla_v h_\epsilon|^2 \rangle$ :

$$\begin{aligned} \frac{d}{dt} \mathcal{H}_\beta(h_\epsilon | 1) &= \frac{2-\beta}{1-\beta} (h^{1-\beta}, h_t) = -\frac{2-\beta}{\epsilon^2} (h_\epsilon^{-\beta} \nabla_v h_\epsilon, \langle h_\epsilon \rangle^\beta \nabla_v h_\epsilon) \\ &\leq -\frac{2-\beta}{\epsilon^2} \|\langle \nabla_v h_\epsilon \rangle\|_{L_x^2}^2 = -\frac{2-\beta}{\epsilon^2} \|\langle v h_\epsilon \rangle\|_{L_x^2}^2. \end{aligned} \tag{5-25}$$

In view of the limiting equation (1-3), we have

$$\frac{d}{dt} \mathcal{H}_\beta(\rho | 1) = \frac{2-\beta}{1-\beta} (\rho^{1-\beta}, \partial_t \rho) = -(2-\beta)(\rho^{-\beta} \nabla_x \rho, \rho^{-\beta} \nabla_x \rho). \tag{5-26}$$

A direct computation with the macroscopic equation (5-3) and (1-3) leads to

$$\begin{aligned} \frac{d}{dt} (\langle h_\epsilon \rangle - \rho, \rho^{1-\beta} - 1) &= (\rho^{1-\beta}, \partial_t \langle h_\epsilon \rangle) + ((1-\beta)\rho^{-\beta} \langle h_\epsilon \rangle - (2-\beta)\rho^{1-\beta}, \partial_t \rho) \\ &= \frac{1-\beta}{\epsilon} (\rho^{-\beta} \nabla_x \rho, \langle v h_\epsilon \rangle) - ((1-\beta)\nabla_x(\rho^{-\beta} \langle h_\epsilon \rangle) - (2-\beta)\rho^{-\beta} \nabla_x \rho, \rho^{-\beta} \nabla_x \rho). \end{aligned} \tag{5-27}$$

The evolution of  $\mathcal{H}_\beta(h_\epsilon | \rho)$  is then estimated by combining (5-25), (5-26), and (5-27):

$$\begin{aligned} \frac{1}{2-\beta} \frac{d}{dt} \mathcal{H}_\beta(h_\epsilon | \rho) &\leq -\frac{1}{\epsilon^2} \|\langle v h_\epsilon \rangle\|_{L_x^2}^2 - \frac{1}{\epsilon} (\langle v h_\epsilon \rangle, \rho^{-\beta} \nabla_x \rho) + (\nabla_x (\rho^{-\beta} \langle h_\epsilon \rangle - \rho^{1-\beta}), \rho^{-\beta} \nabla_x \rho) \\ &= -\|\epsilon^{-1} \langle v h_\epsilon \rangle + \rho^{-\beta} \nabla_x \rho\|_{L_x^2}^2 + (\epsilon^{-1} \langle v h_\epsilon \rangle, \rho^{-\beta} \nabla_x \rho) \\ &\quad + (\nabla_x \langle h_\epsilon \rangle + \beta(1 - \rho^{-1} \langle h_\epsilon \rangle) \nabla_x \rho, \rho^{-2\beta} \nabla_x \rho) \end{aligned}$$

We remark that the above inequality also holds for  $\beta = 1$  by a similar computation. Abbreviate

$$Q_\epsilon := \epsilon^{-1} \langle v h_\epsilon \rangle + \rho^{-\beta} \nabla_x \rho \quad \text{and} \quad R_\epsilon := -\nabla_x \cdot \langle v \otimes \nabla_v h_\epsilon \rangle - \epsilon \partial_t \langle v h_\epsilon \rangle,$$

and write the macroscopic equation (5-4) in the form  $\nabla_x \langle h_\epsilon \rangle = -\epsilon^{-1} \langle h_\epsilon \rangle^\beta \langle v h_\epsilon \rangle + R_\epsilon$ . We then have

$$\begin{aligned} \frac{1}{2-\beta} \frac{d}{dt} \mathcal{H}_\beta(h_\epsilon | \rho) &\leq -\|Q_\epsilon\|_{L_x^2}^2 + ((1 - \rho^{-\beta} \langle h_\epsilon \rangle^\beta) Q_\epsilon, \rho^{-\beta} \nabla_x \rho) + (R_\epsilon, \rho^{-2\beta} \nabla_x \rho) \\ &\quad + (\rho^{-\beta} \langle h_\epsilon \rangle^\beta - \beta \rho^{-1} \langle h_\epsilon \rangle - 1 + \beta, \rho^{-2\beta} |\nabla_x \rho|^2) \\ &\leq 2\|\rho^{-1-\beta} |\nabla_x \rho|\|_{L_{t,x}^\infty}^2 \|\langle h_\epsilon \rangle - \rho\|_{L_x^2}^2 + (R_\epsilon, \rho^{-2\beta} \nabla_x \rho), \end{aligned} \quad (5-28)$$

where, for the second inequality, we used the Cauchy–Schwarz inequality

$$2((1 - \rho^{-\beta} \langle h_\epsilon \rangle^\beta) Q_\epsilon, \rho^{-\beta} \nabla_x \rho) \leq \|Q_\epsilon\|_{L_x^2}^2 + (|1 - \rho^{-\beta} \langle h_\epsilon \rangle^\beta|^2, |\rho^{-\beta} \nabla_x \rho|^2)$$

and the following two elementary inequalities with  $\beta \in [0, 1]$ :

$$|z^\beta - 1| \leq |z - 1| \quad \text{and} \quad |z^\beta - \beta z - 1 + \beta| \leq |z - 1|^2 \quad \text{for any } z \in \mathbb{R}_+.$$

In view of Hölder’s inequality and inequality (5-22) given in Lemma 5.4, we know that

$$\|\langle h_\epsilon \rangle - \rho\|_{L_x^2}^2 \leq \|h_\epsilon - \rho\|_{L^2(dm)}^2 \leq 2\Lambda^\beta \mathcal{H}_\beta(h_\epsilon | \rho).$$

Combining this with (5-28) and (5-20), we derive that, for any  $t \in (0, 1]$ ,

$$\frac{d}{dt} \mathcal{H}_\beta(h_\epsilon | \rho) \leq C_\rho t^{\alpha-1} \mathcal{H}_\beta(h_\epsilon | \rho) + 2(R_\epsilon, \rho^{-2\beta} \nabla_x \rho), \quad (5-29)$$

where the constants  $\alpha \in (0, 1)$  and  $C_\rho > 0$  are provided by Lemma 5.3.

We point out that, after integrating in time, the remainder term in (5-29) involving  $R_\epsilon$  is of order  $O(\epsilon)$  due to the control of the entropy production and the regularity of the limiting equation. Indeed,

$$\begin{aligned} \int_0^t (R_\epsilon, \rho^{-2\beta} \nabla_x \rho) &= \int_0^t (\langle v \otimes \nabla_v h_\epsilon \rangle, \nabla_x (\rho^{-2\beta} \nabla_x \rho)) + \epsilon \int_0^t (\langle v h_\epsilon \rangle, \partial_t (\rho^{-2\beta} \nabla_x \rho)) \\ &\quad - \epsilon (\langle v h_\epsilon \rangle, \rho^{-2\beta} \nabla_x \rho)(t) + \epsilon (\langle v h_\epsilon \rangle, \rho^{-2\beta} \nabla_x \rho)(0) \\ &\lesssim (\|\nabla_x (\rho^{-2\beta} \nabla_x \rho)\|_{L_{t,x}^\infty} + \epsilon \|\partial_t (\rho^{-2\beta} \nabla_x \rho)\|_{L_{t,x}^\infty}) \int_0^t \|\langle v h_\epsilon \rangle\|_{L_x^2} \\ &\quad + \epsilon \|\rho^{-2\beta} \nabla_x \rho\|_{L_{t,x}^\infty} \|\langle v h_\epsilon \rangle\|_{L^\infty([0,T]; L_x^2)}. \end{aligned}$$

It then follows from (5-20), (5-25), and the global upper bound of  $h_\epsilon$  (Lemma 4.1) that, for any  $t \in (0, 1]$ ,

$$\begin{aligned} \int_0^t (R_\epsilon, \rho^{-2\beta} \nabla_x \rho) &\leq C_\rho (t^{(\alpha-1)/2} + \epsilon t^{(\alpha-2)/2}) \left( \int_0^t \|\langle v h_\epsilon \rangle\|_{L_x^2}^2 \right)^{1/2} + C_\rho \epsilon t^{(\alpha-1)/2} \\ &\leq C_\rho (\epsilon t^{(\alpha-1)/2} + \epsilon^2 t^{(\alpha-2)/2}) \sup_{s \in [0, t]} \sqrt{\mathcal{H}_\beta(h_\epsilon | 1)(s)} + C_\rho \epsilon t^{(\alpha-1)/2} \\ &\leq C_\rho (\epsilon t^{(\alpha-1)/2} + \epsilon^2 t^{(\alpha-2)/2}) + C_\rho \epsilon t^{(\alpha-1)/2}. \end{aligned}$$

Combining this estimate with (5-29) as well as Grönwall’s inequality, we conclude (5-23). Additionally, we arrive at (5-24) if we apply Lemma 5.3 with (5-21) instead of (5-20) in the above argument.  $\square$

We are now in a position to conclude the global-in-time diffusion asymptotics.

*Proof of Theorem 1.4.* We are going to combine Propositions 5.1 and 5.5 with a delicate analysis on the relative entropy around the initial time to get Theorem 1.4. The analysis is based on the barrier function method. Let us assume the constant  $\alpha \in (0, 1)$  provided by Proposition 5.5.

Step 1. Pointwise estimate. Let us fix  $\delta \in (0, 1)$  and  $(x_1, v_1) \in \mathbb{T}^d \times B_R$  with  $R > 0$  and consider the function

$$\bar{h}(t, x, v) := C_1 t + C_2 (|x - x_1 - \epsilon^{-1} t v|^2 + |v - v_1|^2),$$

where the constants  $C_1, C_2 > 0$  are to be determined. For any  $t \leq \epsilon \delta / (4(1 + R))$ , we have

$$\bar{h} \geq C_2 (|x - x_1|^2 + |v - v_1|^2 - 2\epsilon^{-1} t |x - x_1| |v|) \geq \frac{1}{2} C_2 \delta^2 = \Lambda \quad \text{on } \partial B_\delta(x_1, v_1),$$

where we chose  $C_2 := 2\delta^{-2} \Lambda$ . For any  $(x, v) \in B_\delta(x_1, v_1)$ ,

$$|\langle h \rangle^\beta \mathcal{L}_{OU} \bar{h}| \lesssim |\Delta_v \bar{h}| + |v \cdot \nabla_v \bar{h}| \lesssim \delta^{-2} (1 + R^2) (1 + \epsilon^{-2} t^2).$$

Therefore, for any  $t \leq \epsilon \delta / (4(1 + R))$  and  $(x, v) \in B_\delta(x_1, v_1)$ ,

$$(\partial_t + \epsilon^{-1} v \cdot \nabla_x - \epsilon^{-2} \langle h \rangle^\beta \mathcal{L}_{OU}) \bar{h} \geq C_1 - C_0 \epsilon^{-2} \delta^{-2} (1 + R^2) (1 + \epsilon^{-2} t^2) \geq 0,$$

where the constant  $C_0 > 0$  is universal and we chose  $C_1 := 2C_0 \epsilon^{-2} \delta^{-2} (1 + R^2)$ . Then, the maximum principle implies that, for any  $t \leq \epsilon \delta / (4(1 + R))$  and  $(x, v) \in B_\delta(x_1, v_1)$ ,

$$|h_\epsilon(t, x, v) - h_{\epsilon, \text{in}}(x_1, v_1)| \leq \bar{h}(t, x, v) + \sup_{B_\delta(x_1, v_1)} |h_{\epsilon, \text{in}}(x, v) - h_{\epsilon, \text{in}}(x_1, v_1)|. \tag{5-30}$$

In particular, for any  $t \leq \epsilon \delta / (4(1 + R))$  and  $(x_1, v_1) \in \mathbb{T}^d \times B_R$ ,

$$|h_\epsilon(t, x_1, v_1) - h_{\epsilon, \text{in}}(x_1, v_1)| \lesssim \epsilon^{-2} \delta^{-2} (1 + R^2) t + \|h_{\epsilon, \text{in}}\|_{C^{\alpha_0}(\mathbb{T}^d \times \mathbb{R}^d)} \delta^\alpha. \tag{5-31}$$

As far as the solution  $\rho$  to the limiting equation (1-3) is concerned, using the Hölder estimate (5-19) in Lemma 5.3, we derive that, for any  $t \in \mathbb{R}_+$ ,

$$\|\rho(t) - \rho_{\text{in}}\|_{L^\infty(\mathbb{T}^d)} \lesssim (1 + \|\rho_{\text{in}}\|_{C^{\alpha_0}(\mathbb{T}^d)}) t^\alpha. \tag{5-32}$$

Step 2. Estimate of the relative entropy around the initial time. Let us restrict our attention to the time interval  $[0, \underline{T}]$  with  $\underline{T} \in (0, 1)$  to be determined. Compute the relative entropy  $\mathcal{H}_\beta(h_\epsilon | \rho)$  in terms of initial data as

$$\begin{aligned} \mathcal{H}_\beta(h_\epsilon | \rho) &= \mathcal{H}_\beta(h_{\epsilon, \text{in}} | \rho_{\text{in}}) + \int_{\mathbb{T}^d \times \mathbb{R}^d} [(\varphi_\beta(h_\epsilon) - \varphi_\beta(h_{\epsilon, \text{in}})) - (\varphi_\beta(\rho) - \varphi_\beta(\rho_{\text{in}}))] \, d\mathbf{m} \\ &\quad - \int_{\mathbb{T}^d \times \mathbb{R}^d} [(\varphi'_\beta(\rho) - \varphi'_\beta(\rho_{\text{in}}))(h_{\epsilon, \text{in}} - \rho_{\text{in}}) + \varphi'_\beta(\rho)(h_\epsilon - h_{\epsilon, \text{in}}) - \varphi'_\beta(\rho)(\rho - \rho_{\text{in}})] \, d\mathbf{m}. \end{aligned} \quad (5-33)$$

Consider a truncation in  $v$  for the integrals on the right-hand side. Since  $h_\epsilon$ ,  $h_{\epsilon, \text{in}}$ ,  $\rho$ , and  $\rho_{\text{in}}$  are all bounded from above ([Lemma 4.1](#)), we have

$$\int_{\mathbb{T}^d \times B_R^c} [|\varphi_\beta(h_\epsilon) - \varphi_\beta(h_{\epsilon, \text{in}})| + |\varphi_\beta(\rho) - \varphi_\beta(\rho_{\text{in}})|] \, d\mathbf{m} \lesssim \int_{B_R^c} d\mu \lesssim R^{-4}. \quad (5-34)$$

Observe that for any  $a, b \in (0, \Lambda]$ , there exists  $\xi \in \mathbb{R}_+$  lying between  $a$  and  $b$  such that

$$\varphi_\beta(a^2) - \varphi_\beta(b^2) = 2\xi \varphi'_\beta(\xi^2)(a - b).$$

Meanwhile,  $|\xi \varphi'_\beta(\xi^2)| \lesssim 1$  for any  $\xi \in (0, \Lambda]$ . Thus,

$$|\varphi_\beta(h_\epsilon) - \varphi_\beta(h_{\epsilon, \text{in}})| \lesssim |h_\epsilon - h_{\epsilon, \text{in}}|^{1/2} \quad \text{and} \quad |\varphi_\beta(\rho) - \varphi_\beta(\rho_{\text{in}})| \lesssim |\rho - \rho_{\text{in}}|^{1/2}.$$

Set  $R := \epsilon^{-\eta/4}$ ,  $\delta := \epsilon^{\eta/4}$ , and  $\underline{T} := \frac{1}{8}\epsilon^{2+2\eta}$  for some constant  $\eta \in (0, 1)$  to be determined. In this setting,  $\underline{T} \leq \epsilon\delta/(4(R))$ . It then follows from (5-31), (5-32), and (5-34) that, for any  $t \leq \underline{T}$ ,

$$\begin{aligned} &\int_{\mathbb{T}^d \times \mathbb{R}^d} [|\varphi_\beta(h_\epsilon) - \varphi_\beta(h_{\epsilon, \text{in}})| + |\varphi_\beta(\rho) - \varphi_\beta(\rho_{\text{in}})|] \, d\mathbf{m} \\ &\quad \lesssim R^{-4} + \int_{\mathbb{T}^d \times B_R} [|h_\epsilon - h_{\epsilon, \text{in}}|^{1/2} + |\rho - \rho_{\text{in}}|^{1/2}] \, d\mathbf{m} \\ &\quad \lesssim \epsilon^\eta + \epsilon^{\eta/2} + \|h_{\epsilon, \text{in}}\|_{C^{\alpha_0}(\mathbb{T}^d \times \mathbb{R}^d)}^{1/2} \epsilon^{\alpha\eta/8} + (1 + \|\rho_{\text{in}}\|_{C^{\alpha_0}(\mathbb{T}^d)}^{1/2}) \epsilon^{\alpha+\alpha\eta}. \end{aligned} \quad (5-35)$$

In addition,

$$\|h_\epsilon - h_{\epsilon, \text{in}}\|_{L^1(\mathbb{T}^d \times B_R^c, d\mathbf{m})} \lesssim R^{-4}$$

and  $|\varphi'_\beta| \lesssim 1$  on  $[\lambda, \Lambda]$ . Therefore, combining (5-31) and (5-32) with inequality (5-22) given in [Lemma 5.4](#) yields, for any  $t \leq \underline{T}$ ,

$$\begin{aligned} &\int_{\mathbb{T}^d \times \mathbb{R}^d} [(\varphi'_\beta(\rho) - \varphi'_\beta(\rho_{\text{in}}))(h_{\epsilon, \text{in}} - \rho_{\text{in}}) + |\varphi'_\beta(\rho)(h_\epsilon - h_{\epsilon, \text{in}})| + |\varphi'_\beta(\rho)(\rho - \rho_{\text{in}})|] \, d\mathbf{m} \\ &\quad \lesssim \|h_{\epsilon, \text{in}} - \rho_{\text{in}}\|_{L^2(\mathbb{T}^d \times \mathbb{R}^d, d\mathbf{m})} + \|h_\epsilon - h_{\epsilon, \text{in}}\|_{L^1(\mathbb{T}^d \times \mathbb{R}^d, d\mathbf{m})} + \|\rho - \rho_{\text{in}}\|_{L^1(\mathbb{T}^d)} \\ &\quad \lesssim \mathcal{H}_\beta^{1/2}(h_{\epsilon, \text{in}} | \rho_{\text{in}}) + R^{-4} + \|h_\epsilon - h_{\epsilon, \text{in}}\|_{L^1(\mathbb{T}^d \times B_R, d\mathbf{m})} + \|\rho - \rho_{\text{in}}\|_{L^1(\mathbb{T}^d)} \\ &\quad \lesssim \epsilon'^{1/2} + \epsilon^\eta + \|h_{\epsilon, \text{in}}\|_{C^{\alpha_0}(\mathbb{T}^d \times \mathbb{R}^d)} \epsilon^{\alpha\eta/4} + (1 + \|\rho_{\text{in}}\|_{C^{\alpha_0}(\mathbb{T}^d)}) \epsilon^{2\alpha+2\alpha\eta}. \end{aligned} \quad (5-36)$$

Plugging (5-35) and (5-36) into expression (5-33), we derive, for any  $t \leq \underline{T}$ ,

$$\mathcal{H}_\beta(h_\epsilon | \rho)(t) \lesssim (1 + \|h_{\epsilon, \text{in}}\|_{C^{\alpha_0}(\mathbb{T}^d \times \mathbb{R}^d)} + \|\rho_{\text{in}}\|_{C^{\alpha_0}(\mathbb{T}^d)}) (\epsilon + \epsilon')^{\alpha\eta/8}. \quad (5-37)$$

**Step 3. Conclusion.** Recall that we have chosen  $\underline{T} = \frac{1}{8}\epsilon^{2+2\eta}$ . In view of (5-37) and the estimate (5-23) given in Proposition 5.5, one may optimize in  $\eta$  to get the result. For simplicity, we pick  $\eta := \alpha/(4 - 2\alpha)$  so that  $\underline{T} = \frac{1}{8}\epsilon^{(4-\alpha)/(2-\alpha)}$  and  $\epsilon(\underline{T}^{(\alpha-1)/2} + \epsilon\underline{T}^{(\alpha-2)/2}) \lesssim \epsilon^{\alpha/2}$ . It turns out that, for any  $t \in [0, 1]$ ,

$$\mathcal{H}_\beta(h_\epsilon | \rho)(t) \leq C_\rho \mathcal{H}_\beta(h_\epsilon | \rho)(\underline{T}) + C_\rho \epsilon (\underline{T}^{(\alpha-1)/2} + \epsilon \underline{T}^{(\alpha-2)/2}) \leq C_* (\epsilon + \epsilon')^{\alpha\eta/8}, \quad (5-38)$$

where the constant  $C_\rho > 0$  is provided in Proposition 5.5 and the constant  $C_* > 0$  depends only on universal constants,  $\|\rho_{\text{in}}\|_{C^{\alpha_0}(\mathbb{T}^d)}$ , and  $\|h_{\epsilon, \text{in}}\|_{C^{\alpha_0}(\mathbb{T}^d \times \mathbb{R}^d)}$ . Then, using the estimate (5-24) given in Proposition 5.5 with (5-22), we arrive at point (i) of Theorem 1.4.

As for point (ii) of Theorem 1.4, applying (5-24) together with (5-22) and (5-38), for any  $t \in [1, \bar{T}]$ , we have

$$\begin{aligned} \|h_\epsilon(t) - \rho(t)\|_{L^2(\text{dm})}^2 &\lesssim [\mathcal{H}_\beta(h_\epsilon | \rho)(1) + \epsilon(1 + \bar{T}^{1/2})]e^{C\bar{T}} \\ &\lesssim [C_* (\epsilon + \epsilon')^{\alpha\eta/8} + \epsilon(1 + (-\iota \log(\epsilon + \epsilon'))^{1/2})](\epsilon + \epsilon')^{-\alpha\eta/16} \lesssim C_* (\epsilon + \epsilon')^{\alpha\eta/16}, \end{aligned}$$

where we picked  $\bar{T} := -\iota \log(\epsilon + \epsilon')$  with  $\iota := \alpha\eta/(16C)$ . Finally, using Proposition 5.1 with the additional assumption that  $h_{\epsilon, \text{in}} \geq \lambda$ , we know from the long-time behavior that there is some universal constant  $c > 0$  such that, for any  $t \geq \bar{T}$ ,

$$\|h_\epsilon(t) - \rho(t)\|_{L^2(\text{dm})} \leq \|h_\epsilon(t) - M_0\|_{L^2(\text{dm})} + \|\rho(t) - M_0\|_{L_x^2} \lesssim e^{-c\bar{T}} = (\epsilon + \epsilon')^{c\iota}. \quad \square$$

## Appendix A: Maximum principle

The following maximum principle (on a not necessarily bounded domain) is repeatedly applied throughout the article. We state it in a more suitable fashion for the Fokker–Planck equations of our concern, whose proof is in the same spirit as [Cameron et al. 2018, Lemma A.2].

**Lemma A.1.** *Let the domain  $\omega$  be a subset of  $\mathbb{R}^d \times \mathbb{R}^d$  and the parabolic cylinder  $\omega_T := (0, T] \times \omega$ . If  $f \in \mathcal{C}_{\text{kin}}^2(\omega_T) \cap \mathcal{C}^0(\bar{\omega}_T)$  is a bounded subsolution in the sense that*

$$\mathcal{L}_1 f := (\partial_t + v \cdot \nabla_x) f - \text{tr}(A D_v^2 f) - B \cdot \nabla_v f \leq 0 \quad \text{in } \omega_T, \quad (\text{A-1})$$

with the coefficients  $A(t, x, v), B(t, x, v) \in \mathcal{C}^0(\omega_T)$  satisfying

$$\lambda|\xi|^2 \leq A(t, x, v)\xi \cdot \xi \leq \Lambda|\xi|^2 \quad \text{and} \quad |B(t, x, v) \cdot \xi| \leq \Lambda|v||\xi| \quad \text{for any } \xi \in \mathbb{R}^d, \quad (t, x, v) \in \omega_T,$$

then  $\sup_{\omega_T} f \leq \sup_{\partial_p \omega_T} f$ , where the parabolic boundary  $\partial_p \omega_T$  is defined to be  $[0, T] \times \bar{\omega} - (0, T] \times \omega$ .

*Proof.* If the domain  $\omega$  is bounded, then the result is classical. For general (unbounded)  $\omega$ , we consider the auxiliary functions

$$\phi_1(t, v) := e^{C_1 t} \langle v \rangle^2 \quad \text{and} \quad \phi_2(t, x) := e^{C_2 t} \langle x \rangle^2$$

with  $C_1, C_2 > 0$ . Since  $f$  is bounded, for any  $\varepsilon_1, \varepsilon_2 > 0$ , there exists  $R(\varepsilon_1), R(\varepsilon_2) > 0$  (independent of  $C_1$  and  $C_2$ ) such that  $f - \varepsilon_1 \phi_1 - \varepsilon_2 \phi_2 \leq \sup_{\partial_p \omega_T} f$  in  $\omega_T \cap \{|x| \geq R(\varepsilon_2) \text{ or } |v| \geq R(\varepsilon_1)\}$ .

By choosing  $C_1 = (d + 2)\Lambda$ , we have

$$\mathcal{L}_1 \phi_1 = e^{C_1 t} (C_1 \langle v \rangle^2 - \text{tr}(A) - 2B \cdot v) \geq (C_1 - (d + 2)\Lambda) \langle v \rangle^2 = 0 \quad \text{in } \omega_T.$$

For any  $R_1 \geq R(\varepsilon_1)$ , there exists  $C_2 > 0$  depending only on  $R_1$  such that

$$\mathcal{L}_1 \phi_2 = e^{C_2 t} (C_2 \langle x \rangle^2 + v \cdot x) \geq (C_2 - 1) \langle x \rangle^2 - |v|^2 \geq 0 \quad \text{in } \omega_T \cap \{|v| < R_1\}.$$

Therefore, for any  $R_2 > R(\varepsilon_2)$ , we have that  $f - \varepsilon_1 \phi_1 - \varepsilon_2 \phi_2$  is a subsolution to (3-1) in the bounded domain  $(0, T] \times (\omega \cap (B_{R_2} \times B_{R_1}))$  with the data smaller than  $\sup_{\partial_p \omega_T} f$  on the boundary portion contained in  $\{|x| = R_2 \text{ or } |v| = R_1\}$ . Then, applying the classical maximum principle yields

$$f - \varepsilon_1 \phi_1 - \varepsilon_2 \phi_2 \leq \sup_{\partial_p \omega_T} f \quad \text{in } (0, T] \times (\omega \cap (B_{R_2} \times B_{R_1})).$$

Sending  $R_2 \rightarrow 0$ ,  $\varepsilon_2 \rightarrow 0$ ,  $R_1 \rightarrow 0$ , and  $\varepsilon_1 \rightarrow 0$  in order, we get the conclusion.  $\square$

### Appendix B: Spreading of positivity

This appendix is devoted to the proof of Proposition 4.2. The argument follows the one presented in [Henderson et al. 2020b], and it is based on the combination of Lemmas 4.5 and 4.6.

*Proof of Proposition 4.2.* The proof is split into four steps.

Step 1. Spreading positivity for all velocities for short times. Applying Lemma 4.5 (with  $\tau = 1$ ) yields that there is some universal constant  $c_0 > 0$  such that, for any  $0 \leq t \leq \min\{1, T, c_0 \langle r^{-1} \rangle^{-2} \langle v_0 \rangle^{-2}\}$ ,

$$h(t, x, v) \geq \frac{1}{8} \delta \mathbb{1}_{\{|x-x_0-tv| < r/2, |v-v_0| < r/2\}} \geq \frac{1}{8} \delta \mathbb{1}_{\{|x-x_0-tv_0| < r/4, |v-v_0| < r/4\}}.$$

Let  $r_0 := \min\{1, \frac{1}{16}r\}$  and  $\underline{t} := \frac{1}{2}\underline{T}$ . Then, Lemma 4.6 implies that there exists  $\underline{C}_0 > 0$  depending only on universal constants,  $\underline{T}$ ,  $\delta$ ,  $r$ , and  $v_0$  such that, for any  $0 < \underline{t} \leq t \leq T_0$  with

$$T_0 := \min\{1, T, c_0 \langle r^{-1} \rangle^{-2} \langle v_0 \rangle^{-2}, \frac{1}{4}r_0 \langle v_0 \rangle^{-1}\}$$

and  $v \in \mathbb{R}^d$ , we have

$$h(t, x, v) \geq \underline{C}_0^{-1} e^{-\underline{C}_0 |v-v_0|^4} \mathbb{1}_{\{|x-x_0-tv_0| < 2r_0\}} \geq \underline{C}_0^{-1} e^{-\underline{C}_0 |v-v_0|^4} \mathbb{1}_{\{|x-x_0| < r_0\}}. \quad (\text{B-1})$$

Step 2. Spreading positivity in space for short times. For any fixed  $\bar{t} \in [\underline{t}, T_0]$  and  $\bar{x} \in \mathbb{T}^d$ , we set  $\bar{v} := (\bar{x} - x_0)/(\bar{t} - \underline{t})$ . In view of (B-1), by Lemma 4.5 (with  $\tau = 2(\bar{t} - \underline{t})$ ,  $v_0 = \bar{v}$ ), we deduce that, if  $\bar{t} - \underline{t} \leq c_0 \langle 2(\bar{t} - \underline{t})r_0^{-1} \rangle^{-2} \langle \bar{v} \rangle^{-2}$  and in particular if

$$\bar{t} \leq \underline{t} + \bar{t}_0 \quad \text{with } \bar{t}_0 := \frac{c_0 r_0^2}{4 + r_0^2} \langle \bar{x} - x_0 \rangle^{-2}, \quad (\text{B-2})$$

then there exists  $\delta_0 > 0$  with the same dependence as  $\underline{C}_0$  such that, for any  $t \in [\underline{t}, \bar{t}]$ ,

$$h(t, x, v) \geq \delta_0 \mathbb{1}_{\{|x-x_0-(t-\underline{t})v| < r_0/2, 2(\bar{t}-\underline{t})|v-\bar{v}| < r_0/2\}} \geq \delta_0 \mathbb{1}_{\{|x-x_0-(t-\underline{t})\bar{v}| < r_0/4, |v-\bar{v}| < r_0/4\}}.$$

Then, Lemma 4.6 (with  $v_0 = \bar{v}$ ) implies that, for any  $0 < 2\underline{t} \leq t \leq \underline{t} + \bar{t}_0$  and  $v \in \mathbb{R}^d$ ,

$$h(t, x, v) \geq \underline{C}_1^{-1} e^{-\underline{C}_1 |v|^4} \mathbb{1}_{\{|x-x_0-(t-\underline{t})\bar{v}| < r_0/8\}} \quad (\text{B-3})$$

for some constant  $\underline{C}_1 > 0$  depending only on universal constants,  $\underline{T}$ ,  $\delta$ ,  $r$ ,  $v_0$ , and  $|\bar{x} - x_0|$ . In particular, for any  $0 < 2\underline{t} \leq \bar{t} \leq \underline{t} + \bar{t}_0$  and  $v \in \mathbb{R}^d$ ,

$$h(\bar{t}, \bar{x}, v) \geq \underline{C}_1^{-1} e^{-\underline{C}_1 |v|^4}. \tag{B-4}$$

Step 3. Spreading positivity for any finite time. We observe that the time interval above is restricted (see (B-2)), but it can be removed by applying the lemmas again. Based on the previous step, it suffices to deal with the case  $\bar{t} > \bar{t}_0$ . By a similar proof to (B-3), we derive

$$h(\bar{t}_0, x, v) \geq \delta_1 \mathbb{1}_{\{|x-\bar{x}| < r_0/8, |v| < r_0/8\}}$$

for some constant  $\delta_1 > 0$  with the same dependence as  $\underline{C}_1$ . In view of this data, applying Lemma 4.5 to  $h(\bar{t}_0 + \cdot, \cdot, \cdot)$  (with  $\tau = 1$ ,  $v_0 = 0$ ), we see that, for any  $t \in [\bar{t}_0, \min\{T_0, \bar{t}_0 + T_1\}]$  with  $T_1 := c_0(8/r_0)^{-2}$ ,

$$h(t, x, v) \geq \frac{1}{8} \delta_1 \mathbb{1}_{\{|x-\bar{x}| < r_0/16, |v| < r_0/16\}}.$$

It then follows from Lemma 4.6 that, for any  $t \in [\bar{t}_0 + \underline{t}, \min\{T_0, \bar{t}_0 + T_1\}]$  and  $v \in \mathbb{R}^d$ ,

$$h(t, x, v) \geq \underline{C}_2^{-1} e^{-\underline{C}_2 |v|^4} \mathbb{1}_{\{|x-\bar{x}| < r_0/32\}}$$

for some constant  $\underline{C}_2 > 0$  with the same dependence as  $\underline{C}_1$ .

Combining this with (B-4) as well as recalling that  $\underline{T} = 2\underline{t}$  and the space domain  $\mathbb{T}^d$  is compact, we know that there exists  $\underline{C}_3 > 0$  depending only on universal constants,  $\underline{T}$ ,  $\delta$ ,  $r$ , and  $v_0$  such that, for any  $(t, x, v) \in [\underline{T}, \min\{T_0, T_1\}] \times \mathbb{T}^d \times \mathbb{R}^d$ ,

$$h(t, x, v) \geq \underline{C}_3^{-1} e^{-\underline{C}_3 |v|^4}.$$

Since  $T_0$  and  $T_1$  depend only on universal constants,  $r$ , and  $v_0$ , by applying the above arguments iteratively, we obtain the result for any finite time.

Step 4. Improving the exponential tail. We remark that this step is not necessary for the applications of the lower bound result, but it shows a more precise decay rate as  $|v| \rightarrow \infty$ .

By the previous step, there is some  $\underline{c} > 0$  depending only on universal constants,  $\underline{T}$ ,  $T$ ,  $\delta$ ,  $r$ , and  $v_0$  such that  $h \geq \underline{c}$  in  $[\underline{T}, T] \times \mathbb{T}^d \times B_1$ . Consider the barrier function

$$\underline{h}(t, x, v) := \underline{c} e^{-C_0(t-\underline{T})^{-1}|v|^2} \quad \text{in } [\underline{T}, T] \times \mathbb{T}^d \times B_1^c,$$

where the constant  $C_0 > 1$  is to be determined. By recalling (4-2) and performing a direct computation, we have

$$\begin{aligned} (\partial_t + v \cdot \nabla_x) \underline{h} - \mathcal{R}_h \mathcal{L}_{OU} \underline{h} &= \frac{C_0 \mathcal{R}_h \underline{h}}{(t-\underline{T})^2} (\mathcal{R}_h^{-1} + 2(d - |v|^2)(t-\underline{T}) - 4C_0|v|^2) \\ &\leq \frac{C_0 \mathcal{R}_h \underline{h}}{(t-\underline{T})^2} (\underline{c}^{-\beta} + 2dT - 4C_0) \quad \text{in } (\underline{T}, T] \times \mathbb{T}^d \times B_1^c. \end{aligned}$$

In particular, by choosing  $C_0$  sufficiently large (with the same dependence as  $\underline{c}$ ), we have

$$(\partial_t + v \cdot \nabla_x)(\underline{h} - h) - \mathcal{R}_h \mathcal{L}_{OU}(\underline{h} - h) \leq 0 \quad \text{in } (\underline{T}, T] \times \mathbb{T}^d \times B_1^c.$$

In addition, by its definition,  $h \geq \underline{h}$  on the boundary  $\{t \in [\underline{T}, T], |v| = 1\} \cup \{t = 2\underline{t}, |v| \geq 1\}$ . The maximum principle (Lemma A.1) then implies that  $h \geq \underline{h}$  in  $[\underline{T}, T] \times \mathbb{T}^d \times B_1^c$ . Therefore, we achieve the Gaussian-type lower bound for any  $(t, x, v) \in [2\underline{T}, T] \times \mathbb{T}^d \times \mathbb{R}^d$ .  $\square$

### Appendix C: Gaining regularity of spatial increment

This appendix is devoted to the proof of two technical lemmas for spatial increments involved in the bootstrapping of higher regularity for solutions to (4-1) presented in Section 4C. For the convenience of the reader, we report a brief proof following the lines of [Imbert and Silvestre 2022, Lemma 8.1] with  $s = 1$  and  $\alpha_1 = \beta = 2$ .

**Lemma C.1.** *Let  $\alpha \in (0, 1)$ , and let a bounded continuous function  $g$  be defined in  $Q_4$ . If there exists some constant  $M > 0$  such that, for any  $y \in B_1$ ,*

$$[\delta_y g]_{C_t^0(Q_2)} \leq M \quad \text{and} \quad [\delta_y g]_{C_t^{2+\alpha}(Q_2)} \leq M \|(0, y, 0)\|^2,$$

then there exists some universal constant  $\eta \in (0, 1)$  such that, for any  $y \in B_1$ ,

$$\|\delta_y g\|_{C_t^\eta(Q_1)} \lesssim M \|(0, y, 0)\|^3.$$

*Proof.* Keeping in mind the assumption and Remark 2.5, for fixed  $y \in B_1$ , we consider the polynomial expansion  $p_0$  of  $\delta_y g$  at  $z_0 \in Q_2$  with  $\deg_{\text{kin}}(p_0) = 2$ :

$$p_0(z) = \delta_y g(z_0) + (\partial_t + v_0 \cdot \nabla_x) \delta_y g(z_0) t + \nabla_v \delta_y g(z_0) \cdot v + \frac{1}{2} D_v^2 \delta_y g(z_0) v \cdot v$$

for  $z := (t, x, v) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d$ . For any  $z$  such that  $z_0 \circ z \in Q_4$ , we have

$$|\delta_y g(z_0 \circ z) - p_0(z)| \leq M \|(0, y, 0)\|^2 \|z\|^{2+\alpha}. \quad (\text{C-1})$$

In particular,  $p_0(0, y, 0) = \delta_y g(z_0)$ , so that, for any  $y \in B_1$ ,

$$\begin{aligned} |\delta_{2y} g(z_0) - 2\delta_y g(z_0)| &= |\delta_y g(z_0 \circ (0, y, 0)) - \delta_y g(z_0)| = |\delta_y g(z_0 \circ (0, y, 0)) - p_0(0, y, 0)| \\ &\leq M \|(0, y, 0)\|^{4+\alpha}. \end{aligned}$$

It then follows that, for any  $z_0 \in Q_2$  and for any  $k \in \mathbb{N}$  such that  $z_0 \circ (0, 2^k y, 0) \in Q_4$ ,

$$\begin{aligned} |\delta_y g(z_0) - 2^{-k} \delta_{2^k y} g(z_0)| &\leq \sum_{j=1}^k 2^{-j} |\delta_{2^j y} g(z_0) - 2\delta_{2^{j-1} y} g(z_0)| \\ &\leq M \|(0, y, 0)\|^{4+\alpha} \sum_{j=1}^k 2^{(1+\alpha)j/3} \leq 2M \|(0, y, 0)\|^{4+\alpha} 2^{(1+\alpha)k/3}. \end{aligned} \quad (\text{C-2})$$

Picking  $k \in \mathbb{N}$  such that  $\|2^{k-1}(0, y, 0)\| \leq 1 < \|2^k(0, y, 0)\|$  and using the assumption yields

$$\begin{aligned} |\delta_y g(z_0)| &\leq 2^{-k} |\delta_{2^k y} g(z_0)| + 2M \|(0, y, 0)\|^{4+\alpha} 2^{(1+\alpha)k/3} \\ &\leq \|\delta_{2^k y} g\|_{C_t^0(Q_2)} \|(0, y, 0)\|^3 + 4M \|(0, y, 0)\|^3 \leq 5M \|(0, y, 0)\|^3. \end{aligned} \quad (\text{C-3})$$



It remains to show that there exists some constant  $\eta > 0$  depending only on  $\alpha$  such that

$$|\delta_y g(z_0 \circ z) - \delta_y g(z_0)| \lesssim M \|(0, y, 0)\|^3 \|z\|^\eta. \tag{C-4}$$

By (C-1) and Lemma 2.4, we know that, for any  $z_0 \in Q_1$  and  $z_0 \circ z \in Q_4$ ,

$$\begin{aligned} |\delta_y g(z_0 \circ z) - \delta_y g(z_0)| &\leq (|(\partial_t + v_0 \cdot \nabla_x) \delta_y g(z_0)| + |D_v^2 \delta_y g(z_0)|) \|z\|^2 \\ &\quad + |\nabla_v \delta_y g(z_0)| \|z\| + M \|(0, y, 0)\|^2 \|z\|^{2+\alpha} \\ &\lesssim ([\delta_y g]_{C_t^{2+\alpha}(Q_2)} \|z\| + [\delta_y g]_{C_t^{2+\alpha}(Q_2)}^{1/2} [\delta_y g]_{C_t^0(Q_2)}^{1/2} + [\delta_y g]_{C_t^0(Q_2)}) \|z\| \\ &\quad + M \|(0, y, 0)\|^2 \|z\|^{2+\alpha}. \end{aligned}$$

If  $\|z\| \leq \|(0, y, 0)\|$ , then combining the above expression with the assumption and (C-3) implies (C-4) with  $\eta = \frac{1}{2}$ . In particular, if  $k \in \mathbb{N}$  such that  $\|z\| < \|2^k(0, y, 0)\|$ , then we have

$$2^{-k} |\delta_{2^k y} g(z_0 \circ z) - \delta_{2^k y} g(z_0)| \lesssim 2^{-k} M \|(0, 2^k y, 0)\|^3 \|z\|^\eta = M \|(0, y, 0)\|^3 \|z\|^\eta. \tag{C-5}$$

Now, if  $\|z\| \geq \|(0, y, 0)\|$ , applying (C-2) at points  $z_0$  and  $z_0 \circ z$ , with  $k \in \mathbb{N}$  such that  $\|2^{k-1}(0, y, 0)\| \leq \|z\| < \|2^k(0, y, 0)\|$ , yields

$$|\delta_y g(z_0) - 2^{-k} \delta_{2^k y} g(z_0)| \leq 4M \|(0, y, 0)\|^3 \|z\|^{1+\alpha}, \tag{C-6}$$

$$|\delta_y g(z_0 \circ z) - 2^{-k} \delta_{2^k y} g(z_0 \circ z)| \leq 4M \|(0, y, 0)\|^3 \|z\|^{1+\alpha}. \tag{C-7}$$

Summing up (C-5), (C-6), and (C-7), we arrive at (C-4). □

Following the lines of the above proof and taking into account that  $\|g\|_{C_t^{2+\alpha}(Q_2)} \leq M$ , one is also able to prove the following result.

**Lemma C.2.** *If  $g \in C_t^{2+\alpha}(Q_2)$  with  $\alpha \in (0, 1)$ , then, for any  $y \in B_1$ , we have*

$$\|\delta_y g\|_{C_t^\alpha(Q_1)} \lesssim \|g\|_{C_t^{2+\alpha}(Q_2)} \|(0, y, 0)\|^2.$$

### Acknowledgements

The authors are grateful to Cyril Imbert for suggesting the question, both François Golse and Cyril Imbert for helpful discussions, and the referees for their careful reading and comments. Zhu’s research has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 754362.

### References

[Anceschi and Polidoro 2020] F. Anceschi and S. Polidoro, “A survey on the classical theory for Kolmogorov equation”, *Matematiche (Catania)* **75**:1 (2020), 221–258. MR Zbl

[Anceschi et al. 2019] F. Anceschi, M. Eleuteri, and S. Polidoro, “A geometric statement of the Harnack inequality for a degenerate Kolmogorov equation with rough coefficients”, *Commun. Contemp. Math.* **21**:7 (2019), art. id. 1850057. MR Zbl

[Arnold et al. 2001] A. Arnold, P. Markowich, G. Toscani, and A. Unterreiter, “On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker–Planck type equations”, *Comm. Partial Differential Equations* **26**:1-2 (2001), 43–100. MR Zbl

- [Bardos et al. 1993] C. Bardos, F. Golse, and C. D. Levermore, “Fluid dynamic limits of kinetic equations, II: Convergence proofs for the Boltzmann equation”, *Comm. Pure Appl. Math.* **46**:5 (1993), 667–753. [MR](#) [Zbl](#)
- [Bouin et al. 2020] E. Bouin, J. Dolbeault, S. Mischler, C. Mouhot, and C. Schmeiser, “Hypocoercivity without confinement”, *Pure Appl. Anal.* **2**:2 (2020), 203–232. [MR](#) [Zbl](#)
- [Cameron et al. 2018] S. Cameron, L. Silvestre, and S. Snelson, “Global a priori estimates for the inhomogeneous Landau equation with moderately soft potentials”, *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **35**:3 (2018), 625–642. [MR](#) [Zbl](#)
- [Chandrasekhar 1943] S. Chandrasekhar, “Stochastic problems in physics and astronomy”, *Rev. Modern Phys.* **15** (1943), 1–89. [MR](#) [Zbl](#)
- [Chavanis 2008] P. H. Chavanis, “Nonlinear mean field Fokker–Planck equations: application to the chemotaxis of biological populations”, *Eur. Phys. J. B* **62**:2 (2008), 179–2008. [Zbl](#)
- [Daskalopoulos and Kenig 2007] P. Daskalopoulos and C. E. Kenig, *Degenerate diffusions: initial value problems and local regularity theory*, EMS Tracts in Math. **1**, Eur. Math. Soc., Zürich, 2007. [MR](#) [Zbl](#)
- [Degond and Mas-Gallic 1987] P. Degond and S. Mas-Gallic, “Existence of solutions and diffusion approximation for a model Fokker–Planck equation”, *Transport Theory Statist. Phys.* **16**:4-6 (1987), 589–636. [MR](#) [Zbl](#)
- [Dolbeault and Li 2018] J. Dolbeault and X. Li, “ $\varphi$ -entropies: convexity, coercivity and hypocoercivity for Fokker–Planck and kinetic Fokker–Planck equations”, *Math. Models Methods Appl. Sci.* **28**:13 (2018), 2637–2666. [MR](#) [Zbl](#)
- [Dolbeault et al. 2015] J. Dolbeault, C. Mouhot, and C. Schmeiser, “Hypocoercivity for linear kinetic equations conserving mass”, *Trans. Amer. Math. Soc.* **367**:6 (2015), 3807–3828. [MR](#) [Zbl](#)
- [El Ghani and Masmoudi 2010] N. El Ghani and N. Masmoudi, “Diffusion limit of the Vlasov–Poisson–Fokker–Planck system”, *Commun. Math. Sci.* **8**:2 (2010), 463–479. [MR](#) [Zbl](#)
- [Esposito et al. 2013] R. Esposito, Y. Guo, C. Kim, and R. Marra, “Non-isothermal boundary in the Boltzmann theory and Fourier law”, *Comm. Math. Phys.* **323**:1 (2013), 177–239. [MR](#) [Zbl](#)
- [Gilbarg and Trudinger 2001] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, Springer, 2001. [MR](#) [Zbl](#)
- [Golse et al. 2019] F. Golse, C. Imbert, C. Mouhot, and A. F. Vasseur, “Harnack inequality for kinetic Fokker–Planck equations with rough coefficients and application to the Landau equation”, *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **19**:1 (2019), 253–295. [MR](#) [Zbl](#)
- [Guo 2002] Y. Guo, “The Landau equation in a periodic box”, *Comm. Math. Phys.* **231**:3 (2002), 391–434. [MR](#) [Zbl](#)
- [Henderson and Snelson 2020] C. Henderson and S. Snelson, “ $C^\infty$  smoothing for weak solutions of the inhomogeneous Landau equation”, *Arch. Ration. Mech. Anal.* **236**:1 (2020), 113–143. [MR](#) [Zbl](#)
- [Henderson et al. 2019] C. Henderson, S. Snelson, and A. Tarfulea, “Local existence, lower mass bounds, and a new continuation criterion for the Landau equation”, *J. Differential Equations* **266**:2-3 (2019), 1536–1577. [MR](#) [Zbl](#)
- [Henderson et al. 2020a] C. Henderson, S. Snelson, and A. Tarfulea, “Local solutions of the Landau equation with rough, slowly decaying initial data”, *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **37**:6 (2020), 1345–1377. [MR](#) [Zbl](#)
- [Henderson et al. 2020b] C. Henderson, S. Snelson, and A. Tarfulea, “Self-generating lower bounds and continuation for the Boltzmann equation”, *Calc. Var. Partial Differential Equations* **59**:6 (2020), art. id. 191. [MR](#) [Zbl](#)
- [Hérau 2018] F. Hérau, “Introduction to hypocoercive methods and applications for simple linear inhomogeneous kinetic models”, pp. 119–147 in *Lectures on the analysis of nonlinear partial differential equations, V* (Beijing, 2014–2015), edited by J.-Y. Chemin et al., Morningside Lect. Math. **5**, Int. Press, Somerville, MA, 2018. [MR](#) [Zbl](#)
- [Imbert and Mouhot 2021] C. Imbert and C. Mouhot, “The Schauder estimate in kinetic theory with application to a toy nonlinear model”, *Ann. H. Lebesgue* **4** (2021), 369–405. [MR](#) [Zbl](#)
- [Imbert and Silvestre 2021] C. Imbert and L. Silvestre, “The Schauder estimate for kinetic integral equations”, *Anal. PDE* **14**:1 (2021), 171–204. [MR](#) [Zbl](#)
- [Imbert and Silvestre 2022] C. Imbert and L. E. Silvestre, “Global regularity estimates for the Boltzmann equation without cut-off”, *J. Amer. Math. Soc.* **35**:3 (2022), 625–703. [MR](#) [Zbl](#)
- [Kato 1975] T. Kato, “The Cauchy problem for quasi-linear symmetric hyperbolic systems”, *Arch. Ration. Mech. Anal.* **58**:3 (1975), 181–205. [MR](#) [Zbl](#)

- [Kim et al. 2020] J. Kim, Y. Guo, and H. J. Hwang, “An  $L^2$  to  $L^\infty$  framework for the Landau equation”, *Peking Math. J.* **3**:2 (2020), 131–202. [MR](#) [Zbl](#)
- [Kolmogoroff 1934] A. Kolmogoroff, “Zufällige Bewegungen (zur Theorie der Brownschen Bewegung)”, *Ann. of Math. (2)* **35**:1 (1934), 116–117. [MR](#) [Zbl](#)
- [Liao et al. 2018] J. Liao, Q. Wang, and X. Yang, “Global existence and decay rates of the solutions near Maxwellian for non-linear Fokker–Planck equations”, *J. Stat. Phys.* **173**:1 (2018), 222–241. [MR](#) [Zbl](#)
- [Manfredini 1997] M. Manfredini, “The Dirichlet problem for a class of ultraparabolic equations”, *Adv. Differential Equations* **2**:5 (1997), 831–866. [MR](#) [Zbl](#)
- [Markou 2017] I. Markou, “Hydrodynamic limit for a Fokker–Planck equation with coefficients in Sobolev spaces”, *Netw. Heterog. Media* **12**:4 (2017), 683–705. [MR](#) [Zbl](#)
- [Moser 1964] J. Moser, “A Harnack inequality for parabolic differential equations”, *Comm. Pure Appl. Math.* **17** (1964), 101–134. [MR](#) [Zbl](#)
- [Mouhot 2018] C. Mouhot, “De Giorgi–Nash–Moser and Hörmander theories: new interplays”, pp. 2467–2493 in *Proceedings of the International Congress of Mathematicians, III* (Rio de Janeiro, 2018), edited by B. Sirakov et al., World Sci., Hackensack, NJ, 2018. [MR](#) [Zbl](#)
- [Nash 1958] J. Nash, “Continuity of solutions of parabolic and elliptic equations”, *Amer. J. Math.* **80** (1958), 931–954. [MR](#) [Zbl](#)
- [Saint-Raymond 2009] L. Saint-Raymond, *Hydrodynamic limits of the Boltzmann equation*, Lecture Notes in Math. **1971**, Springer, 2009. [MR](#) [Zbl](#)
- [Tsallis 1988] C. Tsallis, “Possible generalization of Boltzmann–Gibbs statistics”, *J. Stat. Phys.* **52**:1-2 (1988), 479–487. [MR](#) [Zbl](#)
- [Tsallis 2009] C. Tsallis, *Introduction to nonextensive statistical mechanics: approaching a complex world*, Springer, 2009. [MR](#) [Zbl](#)
- [Villani 2002] C. Villani, “A review of mathematical topics in collisional kinetic theory”, pp. 71–305 in *Handbook of mathematical fluid dynamics, I*, edited by S. Friedlander and D. Serre, North-Holland, Amsterdam, 2002. [MR](#) [Zbl](#)
- [Villani 2009] C. Villani, *Hypocoercivity*, Mem. Amer. Math. Soc. **950**, Amer. Math. Soc., Providence, RI, 2009. [MR](#) [Zbl](#)
- [Yau 1991] H.-T. Yau, “Relative entropy and hydrodynamics of Ginzburg–Landau models”, *Lett. Math. Phys.* **22**:1 (1991), 63–80. [MR](#) [Zbl](#)
- [Zhu 2021] Y. Zhu, “Velocity averaging and Hölder regularity for kinetic Fokker–Planck equations with general transport operators and rough coefficients”, *SIAM J. Math. Anal.* **53**:3 (2021), 2746–2775. [MR](#) [Zbl](#)

Received 25 Feb 2021. Revised 25 May 2022. Accepted 11 Jul 2022.

FRANCESCA ANCESCHI: [f.anceschi@staff.univpm.it](mailto:f.anceschi@staff.univpm.it)

*Dipartimento di Ingegneria Industriale e Scienze Matematiche, Università Politecnica delle Marche, Ancona, Italy*

YUZHE ZHU: [yuzhe.zhu@ens.fr](mailto:yuzhe.zhu@ens.fr)

*Département de mathématiques et applications, École normale supérieure - PSL, Paris, France*

# Analysis & PDE

[msp.org/apde](http://msp.org/apde)

## EDITOR-IN-CHIEF

Clément Mouhot Cambridge University, UK  
[c.mouhot@dpmms.cam.ac.uk](mailto:c.mouhot@dpmms.cam.ac.uk)

## BOARD OF EDITORS

Massimiliano Berti	Scuola Intern. Sup. di Studi Avanzati, Italy <a href="mailto:berti@sissa.it">berti@sissa.it</a>	William Minicozzi II	Johns Hopkins University, USA <a href="mailto:minicozz@math.jhu.edu">minicozz@math.jhu.edu</a>
Zbigniew Blocki	Uniwersytet Jagielloński, Poland <a href="mailto:zbigniew.blocki@uj.edu.pl">zbigniew.blocki@uj.edu.pl</a>	Werner Müller	Universität Bonn, Germany <a href="mailto:mueller@math.uni-bonn.de">mueller@math.uni-bonn.de</a>
Charles Fefferman	Princeton University, USA <a href="mailto:cf@math.princeton.edu">cf@math.princeton.edu</a>	Igor Rodnianski	Princeton University, USA <a href="mailto:irod@math.princeton.edu">irod@math.princeton.edu</a>
David Gérard-Varet	Université de Paris, France <a href="mailto:david.gerard-varet@imj-prg.fr">david.gerard-varet@imj-prg.fr</a>	Yum-Tong Siu	Harvard University, USA <a href="mailto:siu@math.harvard.edu">siu@math.harvard.edu</a>
Colin Guillarmou	Université Paris-Saclay, France <a href="mailto:colin.guillarmou@universite-paris-saclay.fr">colin.guillarmou@universite-paris-saclay.fr</a>	Terence Tao	University of California, Los Angeles, USA <a href="mailto:tao@math.ucla.edu">tao@math.ucla.edu</a>
Ursula Hamenstaedt	Universität Bonn, Germany <a href="mailto:ursula@math.uni-bonn.de">ursula@math.uni-bonn.de</a>	Michael E. Taylor	Univ. of North Carolina, Chapel Hill, USA <a href="mailto:met@math.unc.edu">met@math.unc.edu</a>
Peter Hintz	ETH Zurich, Switzerland <a href="mailto:peter.hintz@math.ethz.ch">peter.hintz@math.ethz.ch</a>	Gunther Uhlmann	University of Washington, USA <a href="mailto:gunther@math.washington.edu">gunther@math.washington.edu</a>
Vadim Kaloshin	Institute of Science and Technology, Austria <a href="mailto:vadim.kaloshin@gmail.com">vadim.kaloshin@gmail.com</a>	András Vasy	Stanford University, USA <a href="mailto:andras@math.stanford.edu">andras@math.stanford.edu</a>
Izabella Laba	University of British Columbia, Canada <a href="mailto:ilaba@math.ubc.ca">ilaba@math.ubc.ca</a>	Dan Virgil Voiculescu	University of California, Berkeley, USA <a href="mailto:dvv@math.berkeley.edu">dvv@math.berkeley.edu</a>
Anna L. Mazzucato	Penn State University, USA <a href="mailto:alm24@psu.edu">alm24@psu.edu</a>	Jim Wright	University of Edinburgh, UK <a href="mailto:j.r.wright@ed.ac.uk">j.r.wright@ed.ac.uk</a>
Richard B. Melrose	Massachusetts Inst. of Tech., USA <a href="mailto:rbm@math.mit.edu">rbm@math.mit.edu</a>	Maciej Zworski	University of California, Berkeley, USA <a href="mailto:zworski@math.berkeley.edu">zworski@math.berkeley.edu</a>
Frank Merle	Université de Cergy-Pontoise, France <a href="mailto:merle@ihes.fr">merle@ihes.fr</a>		

## PRODUCTION

[production@msp.org](mailto:production@msp.org)

Silvio Levy, Scientific Editor

---

See inside back cover or [msp.org/apde](http://msp.org/apde) for submission instructions.

---

The subscription price for 2024 is US \$440/year for the electronic version, and \$690/year (+\$65, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscriber address should be sent to MSP.

---

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online.

---

APDE peer review and production are managed by EditFlow<sup>®</sup> from MSP.

PUBLISHED BY

 **mathematical sciences publishers**  
nonprofit scientific publishing

<http://msp.org/>

© 2024 Mathematical Sciences Publishers

# ANALYSIS & PDE

Volume 17 No. 2 2024

---

On a spatially inhomogeneous nonlinear Fokker–Planck equation: Cauchy problem and diffusion asymptotics	379
FRANCESCA ANCESCHI and YUZHE ZHU	
Strichartz inequalities with white noise potential on compact surfaces	421
ANTOINE MOUZARD and IMMANUEL ZACHHUBER	
Curvewise characterizations of minimal upper gradients and the construction of a Sobolev differential	455
SYLVESTER ERIKSSON-BIQUE and ELEFTERIOS SOULTANIS	
Smooth extensions for inertial manifolds of semilinear parabolic equations	499
ANNA KOSTIANKO and SERGEY ZELIK	
Semiclassical eigenvalue estimates under magnetic steps	535
WAFAA ASSAAD, BERNARD HELFFER and AYMAN KACHMAR	
Necessary density conditions for sampling and interpolation in spectral subspaces of elliptic differential operators	587
KARLHEINZ GRÖCHENIG and ANDREAS KLOTZ	
On blowup for the supercritical quadratic wave equation	617
ELEK CSOBO, IRFAN GLOGIĆ and BIRGIT SCHÖRKHUBER	
Arnold’s variational principle and its application to the stability of planar vortices	681
THIERRY GALLAY and VLADIMÍR ŠVERÁK	
Explicit formula of radiation fields of free waves with applications on channel of energy	723
LIANG LI, RUIPENG SHEN and LIJUAN WEI	
On $L^\infty$ estimates for Monge–Ampère and Hessian equations on nef classes	749
BIN GUO, DUONG H. PHONG, FREID TONG and CHUWEN WANG	