A Kinetic Flocking Model with Diffusion

Renjun Duan, Massimo Fornasier
Johann Radon Institute for Computational and Applied Mathematics
Austrian Academy of Sciences
Altenbergerstrasse 69, A-4040 Linz, Austria

Giuseppe Toscani
Department of Mathematics, University of Pavia
via Ferrata 1, 27100 Pavia, Italy

September 10, 2009

Abstract

We study the stability of the equilibrium states and the rate of convergence of solutions towards them for the continuous kinetic version of the Cucker-Smale flocking in presence of diffusion whose strength depends on the density. This kinetic equation describes the collective behavior of an ensemble of organisms, animals or devices which are forced to adapt their velocities according to a certain rule implying a final configuration in which the ensemble flies at the mean velocity of the initial configuration. Our analysis takes advantage both from the fact that the global equilibrium is a Maxwellian distribution function, and, on the contrary to what happens in the Cucker-Smale model [4], the interaction potential is an integrable function. Precise conditions which guarantee polynomial rates of convergence towards the global equilibrium are found.

Keywords: Flocking, large-time convergence rate, linearized equations.

Contents

1 Introduction

1.1 Main results

Description of the collective and interactive motion of multi-agents such as school of fish, flocking of birds or swarm of bacteria became recently a major research topic in population and behavioral biology and ecology [23, 4, 19, 20, 5, 6]. Among them, the phenomenon of flocking can be regarded as a universal behavior of multi-agents systems, where consensus is reached at large times [4]. Both the numerical and theoretical studies of some related mathematical models which describe various self-organized patterns in the collective motion [3, 26, 16], have shown recently an increasing interest.

In this paper we are concerned with a kinetic flocking model in presence of diffusion. Denoting by $f = f(t, x, \xi) \geq 0$ the number density of particles (e.g. flying birds) which have position $x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$ and velocity $\xi = (\xi_1, \xi_2, \cdots, \xi_n) \in \mathbb{R}^n$ at time $t \geq 0$, $n \geq 1$, the evolution of the density is described by the Fokker-Planck type equation

$$\partial_t f + \xi \cdot \nabla_x f + U * \rho_{\xi f} \cdot \nabla_\xi f = U * \rho_f \nabla_\xi \cdot (\nabla_\xi f + \xi f), \tag{1.1}$$

$$f(0, x, \xi) = f_0(x, \xi), \tag{1.2}$$

where

$$\rho_f(t, x) = \int_{\mathbb{R}^n} f(t, x, \xi) d\xi, \quad \rho_{\xi f}(t, x) = \int_{\mathbb{R}^n} \xi f(t, x, \xi) d\xi.$$

Both the interactive potential $U = U(x)$ and the initial data $f_0 = f_0(x, \xi)$ are given. The operator " $*$ " denotes the convolution with respect to spatial variable. Throughout this paper, it is supposed that U is continuous in x with

$$U(x) = U(|x|) \geq 0, \quad \int_{\mathbb{R}^n} U(x) dx = 1. \tag{1.3}$$

Let us briefly present the origin of the model equation (1.1). When there is no diffusion term in (1.1), the equation

$$\partial_t f + \xi \cdot \nabla_x f + U * \rho_{\xi f} \cdot \nabla_\xi f = U * \rho_f \nabla_\xi \cdot (\xi f) \tag{1.4}$$

has been derived and analyzed by Ha-Tadmor [16] as the mean-field limit of the discrete and finite dimensional flocking model considered by Cucker-Smale [4]. Recently, (1.4) was also obtained as the *grazing collision limit* of a Boltzmann equation of Povzner type in [2]. Within this kinetic picture, the velocities of birds are modified through binary interactions, which dissipate energy according to their mutual distance. Consequently, both equations (1.4) and (1.1) describe a system of particles (e.g. birds) which influence each other according to the potential function of their mutual spatial distance, in such a way that the difference between the respective velocities is diminishing. This can be seen also looking at the characteristic equations, which in the Ha-Tadmor model (1.4) read

$$\begin{cases} \dfrac{dX}{dt} = \Xi, \\ \dfrac{d\Xi}{dt} = \displaystyle\iint_{\mathbb{R}^n \times \mathbb{R}^n} U(|x - y|)(\xi - \Xi) f(t, y, \xi) dy d\xi. \end{cases}$$

Equation (1.1) differs from the Ha-Tadmor model (1.4) in two essential features. In (1.1) particles are subject to random fluctuations whose strength depends on the density, which implies that randomness increases as soon as particles are closer to each other. Resorting again to the collisional kinetic picture, in this new model the velocities of birds are modified through binary interactions, in which, in addition to dissipation of energy according to their mutual distance, also random fluctuations of bird velocities are introduced to mimic a more realistic behavior. The additional presence of random terms, which is reasonable from a physical point of view, is responsible of the existence of a global equilibrium configuration of Maxwellian type. This type of interactions induces a substantial difference in the asymptotic behavior of the solution to (1.1), with respect to model (1.4), where all particles tend exponentially fast to move with their global mean velocity whenever the mutual interaction was strong enough at far distance, independently of the initial conditions. This situation is called *unconditional flocking*, and the global equilibrium (due to the intrinsic dissipation) is represented by a Dirac delta function concentrated at the mean velocity. A second main difference between model (1.4) and the present one is that, on the contrary to what happens in the former, here the interaction potential $U(\cdot)$ is integrable. This corresponds to a weak interaction between birds, or, in other words, to a rapid decay of the interaction in terms of the mutual distance. In consequence of this choice, the relaxation towards equilibrium is not universal, but depends on the size of the perturbation. We remark that also this condition can be reasonably justified from a physical point of view, since it reflects the fact that birds mainly adapt their velocity to birds which are close enough to them. We remark however that the unconditional flocking phenomenon observed in the original Cucker and Smale discrete model [4] heavily depends on the fact that the interaction potential is *not integrable*. Otherwise, results can be recovered only for well-prepared initial configurations of birds.

By direct inspection, one can easily check that the global Maxwellian function

$$\mathbf{M} = \mathbf{M}(\xi) = \frac{1}{(2\pi)^{n/2}} \exp\left(-|\xi|^2/2\right), \tag{1.5}$$

is a steady state of (1.1). Notice that $\mathbf{M}$ has zero bulk velocity and unit density and temperature. The main goal of this paper is to study the stability of solutions near $\mathbf{M}$ and the rate of convergence of these solutions towards it for the Cauchy problem (1.1)-(1.2). For this purpose, introduce the perturbation $u = u(t,x,\xi)$ by setting

$$f = \mathbf{M} + \sqrt{\mathbf{M}}u.$$

Then, u satisfies

$$\begin{aligned}
&\partial_t u + \xi\cdot\nabla_x u + U * \rho_{\xi\sqrt{\mathbf{M}}u}\cdot\nabla_\xi u - \frac{1}{2}U * \rho_{\xi\sqrt{\mathbf{M}}u}\cdot\xi u - U*\rho_{\xi\sqrt{\mathbf{M}}u}\cdot\xi\sqrt{\mathbf{M}}\\
&= \frac{1}{\sqrt{\mathbf{M}}}\int U dx \nabla_\xi\cdot\left(\sqrt{\mathbf{M}}\nabla_\xi u + \frac{1}{2}\xi\sqrt{\mathbf{M}}u\right)\\
&\quad + \frac{1}{\sqrt{\mathbf{M}}}U*\rho_{\sqrt{\mathbf{M}}u}\nabla_\xi\cdot\left(\sqrt{\mathbf{M}}\nabla_\xi u + \frac{1}{2}\xi\sqrt{\mathbf{M}}u\right).
\end{aligned}$$

It is straightforward to verify that

$$\frac{1}{\sqrt{\mathbf{M}}}\nabla_\xi\cdot(\sqrt{\mathbf{M}}\nabla_\xi u + \frac{1}{2}\xi\sqrt{\mathbf{M}}u) = \Delta_\xi u + \frac{1}{4}(2n - |\xi|^2)u.$$

Thus, the Cauchy problem (1.1)-(1.2) is reformulated as

$$\partial_t u + \xi\cdot\nabla_x u + U*\rho_{\xi\sqrt{\mathbf{M}}u}\cdot\nabla_\xi u = \mathbf{L}u + \Gamma(u,u), \tag{1.6}$$

$$u(0,x,\xi) = u_0(x,\xi), \tag{1.7}$$

where u_0 takes the form of

$$u_0 = \mathbf{M}^{-1/2}(f_0 - \mathbf{M}),$$

and the linear part $\mathbf{L}u$ and the nonlinear part $\Gamma(u,u)$ are respectively given by

$$\mathbf{L}u = \Delta_\xi u + \frac{1}{4}(2n - |\xi|^2)u + U*\rho_{\xi\sqrt{\mathbf{M}}u}\cdot\xi\sqrt{\mathbf{M}}, \tag{1.8}$$

$$\Gamma(u,u) = U*\rho_{\sqrt{\mathbf{M}}u}[\Delta_\xi u + \frac{1}{4}(2n-|\xi|^2)u] + \frac{1}{2}U*\rho_{\xi\sqrt{\mathbf{M}}u}\cdot\xi u. \tag{1.9}$$

We introduce some notations. For any integer $m \geq 0$, we use $H^m_{x,\xi}$, H^m_x, H^m_ξ to denote the usual Hilbert spaces $H^m(\mathbb{R}^n_x\times\mathbb{R}^n_\xi)$, $H^m(\mathbb{R}^n_x)$, $H^m(\mathbb{R}^n_\xi)$, respectively, where $L^2_{x,\xi}$, L^2_x, L^2_ξ are also used for $m = 0$.

For a Banach space X, $\|\cdot\|_X$ denotes the corresponding norm, while $\|\cdot\|$ always denotes the norm $\|\cdot\|_{L^2_{x,\xi}}$ for simplicity when $X = L^2_{x,\xi}$. We use $\langle\cdot,\cdot\rangle$ to denote the inner product over the Hilbert space L^2_ξ, i.e.

$$\langle g, h\rangle = \int_{\mathbb{R}^n} g(\xi)h(\xi)d\xi, \quad g, h \in L^2_\xi.$$

For $q \geq 1$, we also define

$$Z_q = L^2_\xi(L^q_x) = L^2(\mathbb{R}^n_\xi; L^q(\mathbb{R}^n_x)), \quad \|g\|_{Z_q} = \left(\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}|g(x,\xi)|^q dx\right)^{2/q} d\xi\right)^{1/2}.$$

Let $\nu(\xi) = 1 + |\xi|^2$. Denote $|\cdot|_\nu$ and $\|\cdot\|_\nu$ by

$$|g|^2_\nu = \int_{\mathbb{R}^n} |\nabla_\xi g(\xi)|^2 + \nu(\xi)|g|^2 d\xi, \quad g = g(\xi), \tag{1.10}$$

$$\|g\|^2_\nu = \iint_{\mathbb{R}^n\times\mathbb{R}^n} |\nabla_\xi g(x,\xi)|^2 + \nu(\xi)|g(x,\xi)|^2 d\xi dx d\xi, \ g = g(x,\xi). \tag{1.11}$$

and $\|\cdot\|_U$ by

$$\|\phi\|^2_U = \iint_{\mathbb{R}^n\times\mathbb{R}^n} U(|x-y|)|\phi(x,y)|^2 dxdy, \quad \phi = \phi(x,y). \tag{1.12}$$

Define the linear operator T_Δ by $T_\Delta b(x,y) = b(x) - b(y)$ for $b = b(x)$. For the multiple indices $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ and $\beta = (\beta_1, \beta_2, \cdots, \beta_n)$, we denote

$$\partial^\alpha_x \partial^\beta_\xi = \partial^{\alpha_1}_{x_1}\partial^{\alpha_2}_{x_2}\cdots\partial^{\alpha_n}_{x_n}\partial^{\beta_1}_{\xi_1}\partial^{\beta_2}_{\xi_2}\cdots\partial^{\beta_n}_{\xi_n}.$$

As usual, the length of α is $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$, and $\alpha' \leq \alpha$ means that $\alpha'_i \leq \alpha_i$ for $1 \leq i \leq n$, while $\alpha' < \alpha$ means $\alpha' \leq \alpha$ and $|\alpha'| < |\alpha|$. For simplicity, we also use ∂_i to denote ∂_{x_i} for each $i = 1, 2, \cdots, n$. In addition, C denotes a generic positive (generally large) constant and λ a generic positive (generally small) constant, where both of them may take different values at different places. When necessary, we write C_0, C_1, $\cdots$, λ_0, λ_1, $\cdots$, to distinguish them.

Now, the main results of this paper are stated as follows.

Theorem 1.1. *Let $n \geq 3$ and $N \geq 2[n/2] + 2$, and let* (1.3) *hold. Suppose that $f_0 \equiv \mathbf{M} + \sqrt{\mathbf{M}}u_0 \geq 0$, and $\|u_0\|_{H^N_{x,\xi}}$ is small enough. Then, the Cauchy problem* (1.6)-(1.7) *admits a unique global solution $u(t,x,\xi)$, satisfying*

$$u \in C([0,\infty); H^N(\mathbb{R}^n\times\mathbb{R}^n)), \quad f \equiv \mathbf{M} + \sqrt{\mathbf{M}}u \geq 0, \tag{1.13}$$

and

$$\|u(t)\|^2_{H^N_{x,\xi}} + \lambda \sum_{|\alpha|+|\beta|\leq N}\int_0^t \|\partial^\alpha_x\partial^\beta_\xi\{\mathbf{I}-\mathbf{P}\}u(s)\|^2_\nu ds + \lambda\sum_{|\alpha|\leq N}\int_0^t \|T_\Delta\partial^\alpha_x b^u(s)\|^2_U ds$$
$$+\lambda\int_0^t \|\nabla_x(a^u, b^u)(s)\|^2_{H^{N-1}_x} ds \leq C\|u_0\|^2_{H^N_{x,\xi}}, \tag{1.14}$$

for any $t \geq 0$, where $\mathbf{P}$, $\mathbf{I}-\mathbf{P}$, a^u, b^u are defined in (2.10). *Moreover, if $\|u_0\|_{Z_1}$ is bounded and $\|u_0\|_{H^N_{x,\xi}} + \|\xi\nabla_x u_0\|$ is small enough, then the time-decay estimate*

$$\|u(t)\|_{H^N_{x,\xi}} \leq C\left(\|u_0\|_{H^N_{x,\xi}} + \|u_0\|_{Z_1}\right)(1+t)^{-\frac{n}{4}}, \tag{1.15}$$

is valid for any $t \geq 0$.

It has to be outlined that (1.1) is a nonlinear Fokker-Planck equation where both the nonlocal drift term and the diffusion coefficient depend on the macroscopic momentum and density, respectively. This kind of nonlinear character leads to the fact that (1.1) does not have the same properties of the classical linear Fokker-Planck equation

$$\partial_t f + \xi \cdot \nabla_x f + \nabla_x V \cdot \nabla_\xi f = \nabla_\xi \cdot (\nabla_\xi f + \xi f), \tag{1.16}$$

where $V = V(x)$ is a confining force potential. In fact, whether or not V is present, (1.16) possesses only one total conservation law (the conservation of mass), while (1.1) conserves not only the total mass but also the total momentum. This difference would imply that the kinetic dissipation of (1.1) should be much weaker than that of (1.16). In addition, (1.16) has a natural Lyapunov functional

$$E_{FP}(f) = \iint_{\mathbb{R}^n \times \mathbb{R}^n} \left[V(x) + \frac{|\xi|^2}{2} + \log f \right] f dx d\xi,$$

nonincreasing in time

$$\frac{d}{dt} E_{FP}(f) = -D_{FP}(f) \equiv - \iint_{\mathbb{R}^n \times \mathbb{R}^n} \frac{1}{f} |\nabla_\xi f + \xi f|^2 dx d\xi \leq 0.$$

In the current case, denoting

$$E(f) = \iint_{\mathbb{R}^n \times \mathbb{R}^n} \left[\frac{|\xi|^2}{2} + \log f \right] f dx d\xi, \tag{1.17}$$

$$D(f) = \iint_{\mathbb{R}^n \times \mathbb{R}^n} \frac{U * \rho_f}{f} |\nabla_\xi f + \xi f|^2 dx d\xi - \int_{\mathbb{R}^n} U * \rho_{\xi f} \cdot \rho_{\xi f} dx, \tag{1.18}$$

solutions to (1.1) satisfy a similar equation

$$\frac{d}{dt} E(f) = -D(f), \tag{1.19}$$

but it is presently unknown if $D(f)$ is non-negative and consequently $E(f)$ is decreasing in time. The eventual existence of a Lyapunov functional for (1.1) is an interesting problem to study. In the case without diffusion Ha-Liu [15] explicitly constructed such Lyapunov functional for (1.4), and used its decay to give a simple proof of the exponential convergence of solutions to the flocking state. Since the nonlinear equation (1.1) lacks such natural a priori bound (only total conservations of mass and momentum hold), we need to turn to the perturbation theory of equilibrium (cf. Theorem 1.1 of this paper) to recover convergence to equilibrium.

We remark that it is straightforward to check that the functional $D(\cdot)$ denoted by (1.18) satisfies

$$D(\mathbf{M}) = 0, \quad \frac{d}{d\epsilon} D(\mathbf{M} + \epsilon\phi)|_{\epsilon=0} = 0, \quad \frac{d^2}{d\epsilon^2} D(\mathbf{M} + \epsilon\phi)|_{\epsilon=0} = 2\mathbf{L}(\frac{\phi}{\sqrt{\mathbf{M}}}).$$

Consequently $\mathbf{M}$ is a critical point of the nonlinear functional $D(\cdot)$. This makes it possible to apply the perturbation method to obtain the stability of equilibrium state $\mathbf{M}$ if the linearized operator $\mathbf{L}$ satisfies certain coercivity inequalities. As stated in Theorem 2.1, it turns out that $\mathbf{L}$ is degenerately dissipative over the full phase space $\mathbb{R}^n_x \times \mathbb{R}^n_\xi$ in the sense of (2.9) below. Then, the classical energy method together with suitable smallness assumptions produce some uniform a priori estimates in high-order Sobolev spaces, which together with the local existence and the continuum argument yield the global existence.

The rate of convergence of solutions to the steady state $\mathbf{M}$ is the other issue under consideration in this paper. For the classical linear Fokker-Planck equation (1.16), thank to the existence of Lyapunov functionals, the hypocoercivity with almost exponential rate or exponential rate in time has been extensively studied by Desvillettes-Villani [7], Mouhot-Neumann [18], Dolbeault-Mouhot-Schmeiser [8] and Villani [25] in a general framework. In the case without diffusion, Ha-Tadmor [16] showed that the energy of solution to (1.4) tends exponentially fast in time to zero for certain strong potential function U. This result has been recently improved in [2], where it has been shown that both the discrete model by Cucker and Smale [4] and its kinetic version (1.4) produce a flocking behavior under the same conditions on the interaction potential.

In the case considered in this paper, as shown in Theorem 1.1, solutions to the Cauchy problem (1.1)-(1.2) which are near equilibrium $\mathbf{M}$, converge to it with an explicit algebraic rate

$$\left\| \frac{f(t)-\mathbf{M}}{\sqrt{\mathbf{M}}} \right\| \leq C_{f_0}(1+t)^{-\frac{n}{4}},$$

for any $t \geq 0$, where C_{f_0} is a constant depending on the size of initial data. As pointed out in [25], the hypocoercivity, which produces the trend towards the equilibrium, essentially stems from the interplay between the conservative free transport operator and the kinetic relaxation. Here, only the algebraic rate is found because the particles move in the whole space $\mathbb{R}^n$ and also the number of the total conservation laws exceed two.

Another issue of this paper is concerned with the direct velocity regularized equation

$$\partial_t f + \xi \cdot \nabla_x f + U * \rho_{\xi f} \cdot \nabla_\xi f = U * \rho_f \nabla_\xi \cdot (\xi f) + \kappa \Delta_\xi f, \tag{1.20}$$

for $\kappa > 0$. Notice that in the above equation, the strength of noise is spatially homogeneous, and the steady state is

$$\mathbf{M}_\kappa = \frac{1}{(2\pi\kappa)^{n/2}} e^{-\frac{|\xi|^2}{2\kappa}}.$$

However, for fixed $\kappa > 0$, it is unclear whether or not $\mathbf{M}_\kappa$ is uniformly stable in time under a certain topology, again due to the lack of a Lyapunov functional. Moreover, the stability of $\mathbf{M}_\kappa$ with $\kappa > 0$ is unknown even for small smooth perturbation of the type considered in Theorem 1.1, because the linearized operator of (1.20) has no coercivity properties similar to that of $\mathbf{L}$ defined in (1.8).

Therefore, as far as equation (1.1) is concerned, it is fundamental for the stability of equilibrium that the strength of noise would depend non-locally on the density. As mentioned before, this dependence implies that randomness is weaker at position x around which the density is lower. A similar phenomenon has been observed in a recent paper [26], where it is argued that coherence in collective swarm motion is facilitated in presence of randomness, which has to be weaker at some position around which mean velocity of particles is larger.

Last, we discuss a variant of the model (1.1). When the potential function U reduces to Dirac delta function concentrated on the origin, the nonlocal nonlinear Fokker-Planck equation (1.1) takes the form

$$\partial_t f + \xi \cdot \nabla_x f = \rho_f \nabla_\xi \cdot [\nabla_\xi f + (\xi - \frac{\rho_{\xi f}}{\rho_f}) f]. \tag{1.21}$$

We refer to [24] for an exhaustive presentation and discussion on the above local nonlinear Fokker-Planck equation. We emphasize that various results including Theorem 1.1 and Theorem 3.1 respectively in the nonlinear and linear cases also hold for equation (1.21). Moreover, when U reduces to Dirac delta function concentrated on the origin, (1.19) remains true for (1.21) with $E(f)$ in (1.17) unchanged and $D(f)$ in (1.18) reducing to $D_0(f)$ given by

$$\begin{aligned} D_0(f) &= \iint_{\mathbb{R}^n \times \mathbb{R}^n} \frac{\rho_f}{f} |\nabla_\xi f + \xi f|^2 dx d\xi - \int_{\mathbb{R}^n} |\rho_{\xi f}|^2 dx \\ &= \iint_{\mathbb{R}^n \times \mathbb{R}^n} \frac{\rho_f}{f} |\nabla_\xi f + (\xi - \frac{\rho_{\xi f}}{\rho_f}) f|^2 dx d\xi \geq 0. \end{aligned}$$

Therefore, (1.21) possesses a natural Lyapunov functional, and the non-perturbation theory would be also possible for the study of well-posedness and large-time behavior of (1.21). This will be object of a separate forthcoming paper.

Finally, we should point out that the present study presents analogies with a recent paper by Guo [14], where the global well-posedness on the torus for the classical Landau equation in absence of external forcing

$$\partial_t f + \xi \cdot \nabla_x f = \nabla_\xi \cdot \left\{ \int_{\mathbb{R}^n} \Lambda(\xi - \xi')[f(\xi') \nabla_\xi f(\xi) - f(\xi) \nabla_\xi f(\xi')] d\xi' \right\}, \tag{1.22}$$

was studied. In Landau equation (1.22) Λ is the non-negative matrix given by

$$\Lambda(\xi) = \Lambda_0 \left(\delta_{ij} - \frac{\xi_i \xi_j}{|\xi|^2} \right) |\xi|^{\gamma+2}, \quad \gamma \geq -3, \ \Lambda_0 > 0.$$

We remark that the main difficulties in this paper rely both in reckoning the kinetic dissipation of the nonlocal linearized operator $\mathbf{L}$ defined in (1.8) and in the control of the nonlinear term in the process of energy estimates. The method used to prove Theorem 3.1, which gives the time-decay estimates on the linearized solution operator is general enough to deal with the time-decay estimates of some other kinetic equations with both the free transport operator and the kinetic relaxation in the full space $\mathbb{R}^n$, whenever the classical spectral analysis is difficult to apply [21]. Actually, we can resort to a similar strategy to obtain results analogous to those of Theorem 3.1 for the Landau equation (1.22), linearized in the case of hard potentials $\gamma \geq 0$, where the space domain is the whole $\mathbb{R}^n$.

The rest of this paper is organized as follows. We shall end this introduction with the next subsection by presenting how equation (1.2) can be derived as a kinetic version of a particle system. In particular, we formally show that this equation arises naturally either as mean-field limit or as grazing collision limit of a Boltzmann type equation from the Cucker-Smale particle model of flocking in presence of an additional stochastic term. In Section 2, for the later study of both the linearized and the original nonlinear equation (1.6), we make two preparations, one of which is to obtain the coercivity of the linearized operator $\mathbf{L}$ as in Theorem 2.1, and the other one to make a macro-micro decomposition of the perturbation u and equation (1.6) where a system of equations for the evolution of moments of u up to second order is derived (cf. (2.21)-(2.24)). In Section 3 we employ Fourier analysis methods to establish the hypocoercivity property for the linearized Cauchy problem with a non-homogeneous microscopic source, and we obtain the precise algebraic time-decay rates. The main idea here is to construct a temporal-frequency free energy functional defined in (3.11) able to capture the macroscopic dissipation in the Fourier space.

In Section 4 we devote ourselves to the proof of the main result for the fully nonlinear Cauchy problem (1.6)-(1.7) (Theorem 1.1 introduced before). To this extent, in Subsection 4.1 we list a series of uniform a priori estimates on the solution, whose proofs are postponed to appendices A.1 and A.2 for the sake of a simpler presentation. The dissipative property of $\mathbf{L}$ proven in Theorem 2.1 and the construction of the other temporal free energy in the phase space contained in (4.6) play a key role in the proof of those a priori estimates. We continue in Subsection 4.2 the proof of the local existence and uniqueness stated in Theorem 4.1 by using an iterative scheme and the standard stability method, concluding with the global existence, which follows by combining the established uniform a priori estimates with the continuum argument. In the last Subsection 4.3 we apply the time-decay properties of the linearized solution operator of Theorem 3.1 to obtain the optimal time-decay rates for the perturbation solution u in some smooth Sobolev space. The main idea of proof is here based on the recently developed energy-spectrum method [12, 11].

1.2 Formal derivation of diffusive model

In this subsection, we shall give a formal derivation of the kinetic equation (1.1). There are at least two ways to do it which we will introduce in what follows. The first way to derive (1.1) is based on the discrete Cucker-Smale model with noise whose strength depends on the distance between particles. Consider the evolutions of m ($m \geq 1$) particles (e.g. birds) with positions and velocities $(x_i, \xi_i) = (x_i(t), \xi_i(t))$ ($1 \leq i \leq m$) at time t in the phase space $\mathbb{R}^n \times \mathbb{R}^n$:

$$\begin{cases} dx_i = \xi_i dt, \\ d\xi_i = \displaystyle\sum_{j=1}^{m} U(|x_j - x_i|)(\xi_j - \xi_i)dt + \sqrt{2\mu \sum_{j=1}^{m} U(|x_j - x_i|) dW_i}. \end{cases} \tag{1.23}$$

Here, U denotes the distance potential (communication rate) function defined as in (1.3). A typical example goes back to the original Cucker-Smale model [4], where

$$U(x) = \frac{C_{n,\gamma}}{(1 + |x|^2)^{\gamma}}, \quad x \in \mathbb{R}^n.$$

In the random noise term, $W_i = W_i(t)$ ($1 \leq i \leq m$) are m independent Wiener processes with values in $\mathbb{R}^n$, and $\mu \geq 0$ is a constant denoting the coefficient of noise strength. Notice that the strength of noise for i-th particle is

$$\mu \sum_{j=1}^{m} U(|x_j - x_i|)$$

which is proportional to the summation of distance potentials of i-th particle with all particles. We remark that if there is only one particle, i.e. $m = 1$, then the system reduces to

$$\begin{cases} dx = \xi dt, \\ d\xi = \sqrt{2\mu U(0)} dW, \end{cases}$$

which means that the motion of a single particle is just a random walk, and further that if $\mu = 0$, then the system is the same as the Cucker-Smale model. We are interested in the so-called *mean field limit* of particle systems (1.23). Thus, set

$$U = \frac{\kappa}{m} U_0$$

for some function U_0 and some constant $\kappa > 0$ which are independent of m, and let

$$f^{(m)}(t, x, \xi) = \frac{1}{m} \sum_{i=1}^{m} \delta(x - x_i(t)) \delta(\xi - \xi_i(t)),$$

where $\delta(\cdot)$ is the Dirac delta function. Since $f^{(m)}$ for each $m \geq 1$ belongs to $\mathcal{M}(\mathbb{R}^{2n})$ which is the space of Radon measure on $\mathbb{R}^{2n}$ and

$$\iint_{\mathbb{R}^n \times \mathbb{R}^n} f^{(m)}(t, x, v) dx dv \equiv 1, \quad t \geq 0,$$

then, up to a subsequence, there is a temporal measure $f(t) \in \mathcal{M}(\mathbb{R}^{2n})$ such that

$$f^{(m)} \to f(t) \text{ in } w^*\text{-}\mathcal{M}(\mathbb{R}^{2n}) \quad \text{as } m \to \infty.$$

Moreover, formally it is a usual way to show (see for instance [15]) that $f(t)$ is the measure-valued weak solution in $\mathcal{M}(\mathbb{R}^{2n})$ to the kinetic equation

$$\partial_t f + \xi \cdot \nabla_x f + \kappa U_0 * \rho_{\xi f} \cdot \nabla_\xi f = \kappa U_0 * \rho_f \nabla_\xi \cdot (\mu \nabla_\xi f + \xi f). \tag{1.24}$$

In equation (1.24) the nonlinear diffusion term follows from the so-called Ito's formula.

On the other hand, the nonlinear kinetic model equation (1.24) of Fokker-Planck type can be also obtained as the grazing limit of a certain kinetic equation of Boltzmann type. Let us assume that the post-interaction velocities (ξ^*, η^*) of two birds which have positions and velocities (x, ξ) and (y, η) before interaction are determined by the law

$$\begin{aligned} \xi^* &= (1 - \kappa U(|x-y|))\xi + \kappa U(|x-y|)\eta + \sqrt{2\mu\kappa U(|x-y|)}\, \theta_\xi, \\ \eta^* &= \kappa U(|x-y|)\xi + (1 - \kappa U(|x-y|))\eta + \sqrt{2\mu\kappa U(|x-y|)}\, \theta_\eta, \end{aligned}$$

where U is defined as before, while $\kappa > 0$ and $\mu \geq 0$ are constants which will enter into the equation exactly in the same way as in (1.24), and

$$\begin{aligned} \theta_\xi &= (\theta_{\xi,1}, \theta_{\xi,2}, \cdots, \theta_{\xi,n}) \in \mathbb{R}^n, \\ \theta_\eta &= (\theta_{\eta,1}, \theta_{\eta,2}, \cdots, \theta_{\eta,n}) \in \mathbb{R}^n. \end{aligned}$$

$\theta_{\xi,i}$ and $\theta_{\eta,i}$, $(1 \leq i \leq n)$ are identically distributed independent random variables of zero mean and unit variance. For this time, it is also supposed that $\sup_x U(|x|)$ is finite and

$$\kappa \sup_x U(|x|) < \frac{1}{2}. \tag{1.25}$$

Notice that this assumption can be removed in the later grazing limit since U will be scaled up to a small parameter $\epsilon > 0$. As in [2], the evolution of the bird density can be described at a kinetic level by the following integro-differential equation of Boltzmann type:

$$\partial_t f + \xi \cdot \nabla_x f = Q(f, f), \tag{1.26}$$

with

$$Q(f,f) = \sigma \iint_{\mathbb{R}^n \times \mathbb{R}^n} \left(\frac{1}{J(|x-y|)} f(x,\xi_*) f(y,\eta_*) - f(x,\xi) f(y,\eta) \right) dy d\eta,$$

where (ξ_*, η_*) mean the pre-collisional velocities of particles that generate the pair velocities (ξ, η) after interaction, and

$$J(|x-y|) = (1 - 2\kappa U(|x-y|))^n$$

is the Jacobian of the transformation of (ξ, η) into (ξ^*, η^*). Notice that J is a well-defined nonnegative function due to the assumption (1.25).

Definition 1.1. *The function $f(t,x,\xi)$ is said to be a weak solution to the Cauchy problem of equation (1.26) with initial data $f_0(x,\xi)$ provided that for any smooth function $\phi(x,\xi)$ with compact support, it holds that*

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^{2n}} \phi(x,\xi) f(t,x,\xi) dx d\xi &= \int_{\mathbb{R}^{2n}} \xi \cdot \nabla_x \phi(x,\xi) f(t,x,\xi) dx d\xi \\ &+ \sigma \mathbf{E} \left[\int_{\mathbb{R}^{4n}} (\phi(x,\xi^*) - \phi(x,\xi)) f(t,x,\xi) f(t,y,\eta) dx dy d\xi d\eta \right] \end{aligned} \tag{1.27}$$

for any $t > 0$ and

$$\lim_{t \to 0+} \int_{\mathbb{R}^{2n}} \phi(x,\xi) f(t,x,\xi) dx d\xi = \int_{\mathbb{R}^{2n}} \phi(x,\xi) f_0(x,\xi) dx d\xi.$$

To carry out the grazing limit, we scale U as

$$U = \epsilon U_0,$$

where $\epsilon > 0$ is a small parameter. Suppose that $f = f(t,x,\xi)$ satisfies the equation (1.26), where f actually takes the form of $f^{\sigma,\epsilon}$ which depends on parameters σ and ϵ but the superscripts are omitted for brevity. Let us begin with the weak form (1.27) and we consider the Taylor's expansion

$$\begin{aligned} \phi(x,\xi^*) - \phi(x,\xi) &= \nabla_\xi \phi(x,\xi) \cdot (\xi^* - \xi) + \frac{1}{2} \sum_{|\beta|=2} \partial_\xi^\beta \phi(x,\xi)(\xi^* - \xi)^\beta \\ &+ \frac{1}{6} \sum_{|\beta|=3} \partial_\xi^\beta \phi(x, \Lambda(\xi^*,\xi))(\xi^* - \xi)^\beta, \end{aligned}$$

where $\Lambda(\xi^*, \xi)$ is a vector between ξ^* and ξ. Recall also that

$$\xi^* - \xi = \epsilon\kappa U_0(|x-y|)(\eta - \xi) + \sqrt{2\mu\epsilon\kappa U_0(|x-y|)}\, \theta_\xi.$$

Then, formally one has

$$\begin{aligned} \mathbf{E}[\phi(x,\xi^*) - \phi(x,\xi)] &= \epsilon \nabla_\xi \phi(x,\xi) \cdot \kappa U_0(|x-y|)(\eta - \xi) \\ &+ \epsilon \Delta_\xi \phi(x,\xi) \cdot \mu\kappa U_0(|x-y|) + O(\epsilon^2). \end{aligned}$$

Thus, it holds

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^{2n}} \phi(x,\xi) f(t,x,\xi) dx d\xi &= \int_{\mathbb{R}^{2n}} \xi \cdot \nabla_x \phi(x,\xi) f(t,x,\xi) dx d\xi \\ &+ \sigma\epsilon \int_{\mathbb{R}^{4n}} \nabla_\xi \phi(x,\xi) \cdot (\eta - \xi) \kappa U_0(|x-y|) f(t,x,\xi) f(t,y,\eta) dx dy d\xi d\eta \\ &+ \sigma\epsilon \int_{\mathbb{R}^{4n}} \mu \Delta_\xi \phi(x,\xi) \kappa U_0(|x-y|) f(t,x,\xi) f(t,y,\eta) dx dy d\xi d\eta \\ &+ \sigma\epsilon O(\epsilon). \end{aligned}$$

Taking the so-called *grazing limit*, so that

$$\epsilon \to 0, \ \ \sigma\epsilon \to 1,$$

then the limit function, still denoted by $f(t, x, \xi)$, satisfies

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{R}^{2n}} \phi(x,\xi) f(t,x,\xi) dx d\xi &= \int_{\mathbb{R}^{2n}} \xi\cdot\nabla_x \phi(x,\xi) f(t,x,\xi) dx d\xi \\ &+ \int_{\mathbb{R}^{4n}} \nabla_\xi \phi(x,\xi)\cdot(\eta-\xi)\kappa U_0(|x-y|) f(t,x,\xi) f(t,y,\eta) dx dy d\xi d\eta \\ &+ \int_{\mathbb{R}^{4n}} \mu \Delta_\xi \phi(x,\xi)\kappa U_0(|x-y|) f(t,x,\xi) f(t,y,\eta) dx dy d\xi d\eta.\end{aligned}$$

This implies that f satisfies

$$\partial_t f + \xi\cdot\nabla_x f = \kappa\nabla_\xi\cdot(f U_0 * (\rho_{\xi f} - \xi\rho_f)) + \kappa\mu\Delta_\xi(f U_0 * \rho_f),$$

which is in the same form as (1.24).

2 Preparations

2.1 Coercivity of the linearized operator

In this subsection, we are concerned with some properties of the linearized operator $\mathbf{L}$ defined by (1.8), especially the coercivity estimate of $\mathbf{L}$ over the Hilbert space $L^2_{x,\xi}$. Notice that $\mathbf{L}$ is the summation of the classical linearized Fokker-Planck operator $\mathbf{L}_{FP}$ and the convolution-type operator $\mathbf{A}$, i.e.

$$\mathbf{L} = \mathbf{L}_{FP} + \mathbf{A}, \tag{2.1}$$

where $\mathbf{L}_{FP}$ and $\mathbf{A}$ are defined by

$$\mathbf{L}_{FP} u = \Delta_\xi u + \frac{1}{4}(2n - |\xi|^2) u, \tag{2.2}$$

$$\mathbf{A} u = U * \rho_{\xi\sqrt{\mathbf{M}}u}\cdot\xi\sqrt{\mathbf{M}}, \tag{2.3}$$

In (2.2) and (2.3) $n \geq 1$ denotes the spatial dimension, U satisfies the condition (1.3) and $\mathbf{M}$ is the normalized global Maxwellian given by (1.5).

Firstly, it is well-known from [1] that $\mathbf{L}_{FP}$ enjoys some dissipative properties, stated in the following

Proposition 2.1. $\mathbf{L}_{FP}$ *is a linear self-adjoint operator with respect to the duality induced by the* L^2_ξ*-scalar product, and it is local in* x*. Furthermore, the following properties hold.*

(i) One has

$$\langle \mathbf{L}_{FP} u, u\rangle = -\int_{\mathbb{R}^n} \left|\nabla_\xi\left(\frac{u}{\sqrt{\mathbf{M}}}\right)\right|^2 \mathbf{M} d\xi,$$

$$\mathrm{Ker}\ \mathbf{L}_{FP} = \mathrm{Span}\{\sqrt{\mathbf{M}}\}, \quad \mathrm{Range}\ \mathbf{L}_{FP} = \mathrm{Span}\{\sqrt{\mathbf{M}}\}^\perp.$$

(ii) Define the projector $\mathbf{P}_0$ *by*

$$\mathbf{P}_0 u = a^u\sqrt{\mathbf{M}}, \quad a^u \equiv \langle\sqrt{\mathbf{M}}, u\rangle.$$

Then, one has the identity

$$\begin{aligned}\langle \mathbf{L}_{FP} u, u\rangle = &-\int_{\mathbb{R}^n} |\nabla_\xi\{\mathbf{I}-\mathbf{P}_0\}u|^2 d\xi - \frac{1}{4}\int_{\mathbb{R}^n} |\xi|^2|\{\mathbf{I}-\mathbf{P}_0\}u|^2 d\xi \\ &+\frac{n}{2}\int_{\mathbb{R}^n} |\{\mathbf{I}-\mathbf{P}_0\}u|^2 d\xi.\end{aligned}$$

(iii) There exists a constant $\lambda_{FP} > 0$ such that the Poincaré inequality holds:

$$-\langle \mathbf{L}_{FP} u, u\rangle \geq \lambda_{FP} \int_{\mathbb{R}^n} |\{\mathbf{I}-\mathbf{P}_0\}u|^2 d\xi.$$

(iv) More strongly, there is a constant $\lambda_0 > 0$ such that the coercivity estimate holds:

$$-\langle \mathbf{L}_{FP} u, u\rangle \geq \lambda_0 |\{\mathbf{I}-\mathbf{P}_0\}u|_\nu^2, \tag{2.4}$$

where the norm $|\cdot|_\nu$ is defined in (1.10).

Next, we shall obtain some coercivity estimate similar to (2.4) on the non-local linear operator $\mathbf{L}$ in the phase space $\mathbb{R}^n_x \times \mathbb{R}^n_\xi$. Notice that it is straightforward to make estimates on $\mathbf{A}$ as

$$\left|\int_{\mathbb{R}^n} \langle \mathbf{A}u, u\rangle dx\right| \leq C_n \iint_{\mathbb{R}^n\times\mathbb{R}^n} |\{\mathbf{I}-\mathbf{P}_0\}u|^2 dx d\xi,$$

where $C_n = \langle |\xi|^2, \mathbf{M}\rangle$ depends only on n. On the other hand, from (2.4) it holds

$$-\int_{\mathbb{R}^n} \langle \mathbf{L}_{FP} u, u\rangle dx \geq \lambda_0 \|\{\mathbf{I}-\mathbf{P}_0\}u\|_\nu^2.$$

Since it is not clear presently whether λ_0 is strictly larger than C_n, it is nontrivial to get a coercivity estimate on $\mathbf{L}$ directly from (2.4). It turns out that one has to extract part of dissipation of $\mathbf{L}_{FP}$ corresponding to the momentum component of u in order to control the non-local operator $\mathbf{A}$. To do that, let us decompose the Hilbert space L^2_ξ as

$$L^2_\xi = \mathcal{N} \oplus \mathcal{N}^\perp, \quad \mathcal{N} = \mathrm{Span}\{\sqrt{\mathbf{M}}, \xi\sqrt{\mathbf{M}}\},$$

and define the projector $\mathbf{P}$ by

$$\mathbf{P} : L^2_\xi \to \mathcal{N}, \quad u \mapsto \mathbf{P}u \equiv \{a^u + b^u\cdot\xi\}\sqrt{\mathbf{M}}.$$

Notice that since $\sqrt{\mathbf{M}}, \xi_1\sqrt{\mathbf{M}}, \cdots, \xi_n\sqrt{\mathbf{M}}$ forms an orthonormal basis of $\mathcal{N}$, then one has

$$a^u = \langle \sqrt{\mathbf{M}}, u\rangle, \quad b^u = \langle \xi\sqrt{\mathbf{M}}, u\rangle.$$

We also introduce the projector $\mathbf{P}_1$ by

$$\mathbf{P}_1 u = b^u\cdot\xi\sqrt{\mathbf{M}} = \langle \xi\sqrt{\mathbf{M}}, u\rangle \cdot \xi\sqrt{\mathbf{M}}.$$

Then $\mathbf{P}$ can be written as

$$\mathbf{P} = \mathbf{P}_0 \oplus \mathbf{P}_1,$$

in L^2_ξ. The main result of this subsection concerning the coercivity estimate of $\mathbf{L}$ is stated as follows.

Theorem 2.1. *Let $n \geq 1$ and* (1.3) *hold. The operators $\mathbf{L}, \mathbf{L}_{FP}, \mathbf{A}$ are defined by* (2.1), (2.2) *and* (2.3), *respectively. Then, the following holds.*

(i) $\mathbf{A}$ and hence $\mathbf{L}$ are linear nonlocal operators which are self-adjoint with respect to the duality induced by the $L^2_{x,\xi}$-scalar product;

(ii) One has identities:

$$\mathbf{A}u = \mathbf{A}\mathbf{P}u = \mathbf{P}\mathbf{A}u = \mathbf{A}\mathbf{P}_1 u = \mathbf{P}_1\mathbf{A}u = U * b^u\cdot\xi\sqrt{\mathbf{M}}, \tag{2.5}$$

$$\mathbf{A}\{\mathbf{I}-\mathbf{P}\}u = \{\mathbf{I}-\mathbf{P}\}\mathbf{A}u = 0, \tag{2.6}$$

$$\mathbf{L}_{FP}\mathbf{P}u = \mathbf{L}_{FP}\mathbf{P}_1 u = -\mathbf{P}_1 u, \ \ \mathbf{P}\mathbf{L}_{FP} u = -\mathbf{P}_1 u, \tag{2.7}$$

$$\mathbf{L}\mathbf{P}u = -[\mathbf{P}_1, \mathbf{A}]u = -(b^u - U * b^u)\cdot\xi\sqrt{\mathbf{M}}, \tag{2.8}$$

where $[\mathbf{P}_1, \mathbf{A}]$ denotes the commutator $\mathbf{P}_1\mathbf{A} - \mathbf{A}\mathbf{P}_1$;

(iii) Let λ_0 be defined in (2.4). *Then, the coercivity inequality*

$$-\int_{\mathbb{R}^n} \langle \mathbf{L}u, u\rangle dx \geq \lambda_0 \|\{\mathbf{I}-\mathbf{P}\}u\|_\nu^2 + \frac{1}{2}\|T_\Delta b^u\|_U^2, \tag{2.9}$$

holds for any $u = u(x,\xi)$, where the norms $\|\cdot\|_\nu$ and $\|\cdot\|_U$ are defined in (1.11) *and* (1.12), *respectively.*

Proof. To prove (i), for any $u = u(x,\xi), v = v(x,\xi)$, it holds

$$\begin{aligned}\int_{\mathbb{R}^n} \langle \mathbf{A}u, v\rangle dx &= \iint_{\mathbb{R}^n\times\mathbb{R}^n} U * b^u \cdot \xi\sqrt{\mathbf{M}} v dx d\xi = \int_{\mathbb{R}^n} U * b^u b^v dx \\ &= \iint_{\mathbb{R}^n\times\mathbb{R}^n} U(|x-y|) b^u(y)\cdot b^v(x) dx dy \\ &= \int_{\mathbb{R}^n} b^u U * b^v dx = \int_{\mathbb{R}^n} \langle u, \mathbf{A}v\rangle dx,\end{aligned}$$

where the symmetry of $U = U(|x|)$ was used. Then, $\mathbf{A}$ is self-adjoint on $L^2_{x,\xi}$. Since $\mathbf{L}_{FP}$ is also self-adjoint with respect to the duality induced by the $L^2_{x,\xi}$-scalar product, so is $\mathbf{L} = \mathbf{L}_{FP} + \mathbf{A}$.

To prove (ii), (2.5) and (2.6) directly follow from definitions of $\mathbf{A}$, $\mathbf{P}_1$ and $\mathbf{P}$. Notice

$$\mathbf{L}_{FP}\mathbf{P}u = \mathbf{L}_{FP}\mathbf{P}_0 u + \mathbf{L}_{FP}\mathbf{P}_1 u = \mathbf{L}_{FP}\mathbf{P}_1 u$$

holds from Proposition 2.1 (i). Then, one can compute

$$\mathbf{L}_{FP}\mathbf{P}u = \Delta_\xi(b^u\cdot\xi\sqrt{\mathbf{M}}) + \frac{1}{4}(2n - |\xi|^2)(b^u\cdot\xi\sqrt{\mathbf{M}}) = -b^u\cdot\xi\sqrt{\mathbf{M}} = -\mathbf{P}_1 u,$$

where we used

$$\Delta_\xi(b^u\cdot\xi\sqrt{\mathbf{M}}) = \Delta_\xi(b^u\cdot\xi) + 2\nabla_\xi(b^u\cdot\xi)\cdot\nabla_\xi\sqrt{\mathbf{M}} + b^u\cdot\xi\Delta_\xi\sqrt{\mathbf{M}},$$

and

$$\nabla_\xi\sqrt{\mathbf{M}} = -\frac{1}{2}\xi\sqrt{\mathbf{M}}, \quad \Delta_\xi\sqrt{\mathbf{M}} = -\frac{n}{2}\sqrt{\mathbf{M}} + \frac{1}{4}|\xi|^2\sqrt{\mathbf{M}}.$$

Moreover, it holds

$$\begin{aligned}\mathbf{P}\mathbf{L}_{FP}u &= \langle\sqrt{\mathbf{M}}, \mathbf{L}_{FP}u\rangle\sqrt{\mathbf{M}} + \langle\xi\sqrt{\mathbf{M}}, \mathbf{L}_{FP}u\rangle\cdot\xi\sqrt{\mathbf{M}} \\ &= \langle\mathbf{L}_{FP}\mathbf{P}_0\sqrt{\mathbf{M}}, u\rangle\sqrt{\mathbf{M}} + \langle\mathbf{L}_{FP}\mathbf{P}_1(\xi\sqrt{\mathbf{M}}), u\rangle\cdot\xi\sqrt{\mathbf{M}} \\ &= -\langle\xi\sqrt{\mathbf{M}}, u\rangle\cdot\xi\sqrt{\mathbf{M}} = -\mathbf{P}_1 u.\end{aligned}$$

Then, equation (2.7) is proved. Equation (2.8) follows from $\mathbf{L} = \mathbf{L}_{FP} + \mathbf{A}$ and (2.6)-(2.7).

To prove (iii), for any u, one has

$$\begin{aligned}\langle\mathbf{L}_{FP}u, u\rangle &= \langle\mathbf{L}_{FP}\mathbf{P}u, \mathbf{P}u\rangle + \langle\mathbf{L}_{FP}\mathbf{P}u, \{\mathbf{I}-\mathbf{P}\}u\rangle \\ &\quad + \langle\mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}u, \mathbf{P}u\rangle + \langle\mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}u, \{\mathbf{I}-\mathbf{P}\}u\rangle \\ &= \langle\mathbf{L}_{FP}\mathbf{P}u, \mathbf{P}u\rangle + 2\langle\mathbf{L}_{FP}\mathbf{P}u, \{\mathbf{I}-\mathbf{P}\}u\rangle \\ &\quad + \langle\mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}u, \{\mathbf{I}-\mathbf{P}\}u\rangle,\end{aligned}$$

where for the first two terms on the r.h.s., further from (2.7), it holds

$$\begin{aligned}\langle\mathbf{L}_{FP}\mathbf{P}u, \mathbf{P}u\rangle &= -\langle\mathbf{P}_1 u, \mathbf{P}u\rangle = -\langle\mathbf{P}_1 u, \mathbf{P}_1 u\rangle = -|b^u|^2, \\ \langle\mathbf{L}_{FP}\mathbf{P}u, \{\mathbf{I}-\mathbf{P}\}u\rangle &= -\langle\mathbf{P}_1 u, \{\mathbf{I}-\mathbf{P}\}u\rangle = 0.\end{aligned}$$

Then, one has

$$\langle\mathbf{L}_{FP}u, u\rangle = \langle\mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}u, \{\mathbf{I}-\mathbf{P}\}u\rangle - |b^u|^2.$$

On the other hand, for $\mathbf{A}$, similarly one has

$$\int_{\mathbb{R}^n}\langle\mathbf{A}u, u\rangle dx = \int_{\mathbb{R}^n}\langle\mathbf{A}\mathbf{P}_1 u, u\rangle dx = \int_{\mathbb{R}^n}\langle\mathbf{P}_1 u, \mathbf{A}\mathbf{P}_1 u\rangle dx = \int_{\mathbb{R}^n} U * b^u\cdot b^u dx.$$

Thus, combining the above estimates on $\mathbf{L}_{FP}$ and $\mathbf{A}$, it follows that

$$\int_{\mathbb{R}^n} \langle \mathbf{L}u, u\rangle dx = \int_{\mathbb{R}^n} \langle \mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}u, \{\mathbf{I}-\mathbf{P}\}u\rangle dx \\ - \int_{\mathbb{R}^n} |b^u|^2 dx + \int_{\mathbb{R}^n} U * b^u \cdot b^u dx.$$

One can further compute

$$\begin{aligned}
&\int_{\mathbb{R}^n} |b^u|^2 dx - \int_{\mathbb{R}^n} U * b^u \cdot b^u dx \\
&= \iint_{\mathbb{R}^n\times\mathbb{R}^n} U(|x-y|) b^u(x)(b^u(x) - b^u(y)) dxdy \\
&= \iint_{\mathbb{R}^n\times\mathbb{R}^n} U(|y-x|) b^u(y)(b^u(y) - b^u(x)) dxdy \\
&= \frac{1}{2}\iint_{\mathbb{R}^n\times\mathbb{R}^n} U(|x-y|) |b^u(x) - b^u(y)|^2 dxdy = \frac{1}{2}\|T_\Delta b^u\|_U^2.
\end{aligned}$$

Therefore, (2.9) follows from the coercivity inequality (2.4) for $\mathbf{L}_{FP}$ and

$$\{\mathbf{I}-\mathbf{P}_0\}\{\mathbf{I}-\mathbf{P}\} = \{\mathbf{I}-\mathbf{P}\}.$$

Hence, also (iii) is proved. This completes the proof of Theorem 2.1. □

2.2 Macro-micro decomposition

As usual, for fixed (t,x), $u(t,x,\xi)$ can be uniquely decomposed as

$$\begin{cases} u(t,x,\xi) = \mathbf{P}u + \{\mathbf{I}-\mathbf{P}\}u, \\ \mathbf{P}u \equiv \{a^u + b^u\cdot\xi\}\sqrt{\mathbf{M}}, \\ a^u = \langle\sqrt{\mathbf{M}}, u\rangle, \quad b^u = \langle \xi\sqrt{\mathbf{M}}, u\rangle, \end{cases} \tag{2.10}$$

where $\mathbf{P}u$ is called the macroscopic component of u while $\{\mathbf{I}-\mathbf{P}\}u$ is called the corresponding microscopic component. Notice that by the definitions of a^u and b^u, it holds

$$\mathbf{P}u \perp \{\mathbf{I}-\mathbf{P}\}u \tag{2.11}$$

in L^2_ξ for any (t,x).

In what follows, let us suppose that u satisfies the perturbation equation (1.6) and the spatial dimension $n \geq 1$ holds. For later use, let us now derive some macroscopic balance laws satisfied by the macro components a^u and b^u. To do that, rewrite (1.6) as

$$\partial_t u + \xi\cdot\nabla_x u + U * b^u \cdot \nabla_\xi u = \mathbf{L}u + \Gamma(u,u), \tag{2.12}$$

where by (1.9), $\Gamma(\cdot,\cdot)$ is regarded as a bilinear operator defined by

$$\Gamma(u,v) = U * a^u \mathbf{L}_{FP} v + \frac{1}{2} U * b^u \cdot \xi v. \tag{2.13}$$

After taking velocity integration from the unperturbed equation (1.1), one has the local conservation law of mass:

$$\partial_t \int_{\mathbb{R}^n} f d\xi + \nabla_x \cdot \int_{\mathbb{R}^n} \xi f d\xi = 0, \tag{2.14}$$

and the local balance law of momentum:

$$\partial_t \int_{\mathbb{R}^n} \xi_i f d\xi + \nabla_x \cdot \int_{\mathbb{R}^n} \xi\xi_i f d\xi - U * \rho_{\xi_i f} \int_{\mathbb{R}^n} f d\xi = -U * \rho_f \int_{\mathbb{R}^n} \xi_i f d\xi. \tag{2.15}$$

for $1 \le i \le n$. By using the macro-micro decomposition (2.10) and the property (2.11), one can compute the moments of f up to second order as follows:

$$\int_{\mathbb{R}^n} f d\xi = \int_{\mathbb{R}^n} (\mathbf{M} + \sqrt{\mathbf{M}} u) d\xi = 1 + a^u,$$
$$\int_{\mathbb{R}^n} \xi_i f d\xi = \int_{\mathbb{R}^n} \xi_i (\mathbf{M} + \sqrt{\mathbf{M}} u) d\xi = b_i^u,$$

and

$$\begin{aligned}\int_{\mathbb{R}^n} \xi_i \xi_j f d\xi &= \int_{\mathbb{R}^n} \xi_i \xi_j (\mathbf{M} + \sqrt{\mathbf{M}} u) d\xi \\ &= \delta_{ij} + \int_{\mathbb{R}^n} \xi_i \xi_j \sqrt{\mathbf{M}} \mathbf{P} u d\xi + \int_{\mathbb{R}^n} \xi_i \xi_j \sqrt{\mathbf{M}} \{\mathbf{I} - \mathbf{P}\} u d\xi \\ &= (1 + a^u) \delta_{ij} + \langle \xi_i \xi_j \sqrt{\mathbf{M}}, \{\mathbf{I} - \mathbf{P}\} u \rangle,\end{aligned}$$

for $1 \le i, j \le n$, where δ_{ij} is the Kronecker delta. Thus, it follows from (2.14) and (2.15) that

$$\partial_t a^u + \nabla_x \cdot b^u = 0,$$

and

$$\begin{aligned}\partial_t b_i^u + \partial_i a^u + \sum_j \partial_j \langle \xi_i \xi_j \sqrt{\mathbf{M}}, \{\mathbf{I} - \mathbf{P}\} u \rangle - U * b_i^u (1 + a^u) \\ = -U * (1 + a^u) b_i^u, \quad 1 \le i \le n.\end{aligned}$$

Next, we need to derive the evolution of second-order moments of $\{\mathbf{I} - \mathbf{P}\} u$:

$$\langle \xi \otimes \xi \sqrt{\mathbf{M}}, \{\mathbf{I} - \mathbf{P}\} u \rangle.$$

By using

$$\mathbf{L} = \mathbf{L}_{FP} + \mathbf{A}, \quad \mathbf{L}_{FP} \mathbf{P} = -\mathbf{P}_1,$$

and the macro-micro decomposition (2.10), one can further rewrite (2.12) as

$$\begin{aligned}\partial_t u + \xi \cdot \nabla_x u + U * b^u \cdot \nabla_\xi u = \mathbf{L}_{FP} \{\mathbf{I} - \mathbf{P}\} u + (U * b^u - b^u) \cdot \xi \sqrt{\mathbf{M}} \\ + U * a^u \mathbf{L}_{FP} \{\mathbf{I} - \mathbf{P}\} u - U * a^u b^u \cdot \xi \sqrt{\mathbf{M}} \\ + \frac{1}{2} U * b^u \cdot \xi \mathbf{P} u + \frac{1}{2} U * b^u \cdot \xi \{\mathbf{I} - \mathbf{P}\} u,\end{aligned}$$

that is

$$\begin{aligned}\partial_t \mathbf{P} u + \xi \cdot \nabla_x \mathbf{P} u + U * b^u \cdot \nabla_\xi \mathbf{P} u \\ -(U * b^u - b^u) \cdot \xi \sqrt{\mathbf{M}} + U * a^u b^u \cdot \xi \sqrt{\mathbf{M}} - \frac{1}{2} U * b^u \cdot \xi \mathbf{P} u \\ = -\partial_t \{\mathbf{I} - \mathbf{P}\} u + l + r,\end{aligned} \tag{2.16}$$

In (2.16), the linear term l and the nonlinear term r, are given respectively by

$$l = -\xi \cdot \nabla_x \{\mathbf{I} - \mathbf{P}\} u + \mathbf{L}_{FP} \{\mathbf{I} - \mathbf{P}\} u, \tag{2.17}$$
$$r = U * a^u \mathbf{L}_{FP} \{\mathbf{I} - \mathbf{P}\} u + \frac{1}{2} U * b^u \cdot \xi \{\mathbf{I} - \mathbf{P}\} u - U * b^u \cdot \nabla_\xi \{\mathbf{I} - \mathbf{P}\} u. \tag{2.18}$$

By using the representation of $\mathbf{P} u$ as in (2.10), one can expand the l.h.s. of (2.16) as

$$\begin{aligned}&\{\partial_t a^u + U * b^u \cdot b^u\} \sqrt{\mathbf{M}} \\ &+ \sum_{i=1}^n \{\partial_t b_i^u + \partial_i a^u - (U * b_i^u - b_i^u) + U * a^u b_i^u - U * b_i^u a^u\} \xi_i \sqrt{\mathbf{M}} \\ &+ \sum_{ij=1}^n \{\partial_i b_j^u - U * b_i^u b_j^u\} \xi_i \xi_j \sqrt{\mathbf{M}} \\ &= -\partial_t \{\mathbf{I} - \mathbf{P}\} u + l + r.\end{aligned} \tag{2.19}$$

Let us define the moment function $A = (A_{ij}(\cdot))_{n\times n}$ by

$$A_{ij}(u) = \int_{\mathbb{R}^n} (\xi_i\xi_j - 1)\sqrt{\mathbf{M}} u d\xi. \tag{2.20}$$

Then, applying $A_{ij}(\cdot)$ to both sides of (2.19) yields

$$\partial_i b_i^u - U * b_i^u b_i^u = -\partial_t A_{ii}(\{\mathbf{I}-\mathbf{P}\}u) + A_{ii}(l+r),$$

and

$$\partial_i b_j^u + \partial_j b_i^u - U * b_i^u b_j^u - U * b_j^u b_i^u = -\partial_t A_{ij}(\{\mathbf{I}-\mathbf{P}\}u) + A_{ij}(l+r),\ i \neq j,$$

where $1 \le i, j \le n$.

In summary, the macro components a^u and b^u satisfy the equations

$$\partial_t a^u + \nabla_x \cdot b^u = 0, \tag{2.21}$$

$$\begin{aligned}\partial_t b_i^u + \partial_i a^u - (U * b_i^u - b_i^u) + U * a^u b_i^u - U * b_i^u a^u \\ + \sum_{j=1}^n \partial_j A_{ij}(\{\mathbf{I}-\mathbf{P}\}u) = 0,\end{aligned} \tag{2.22}$$

$$\partial_t A_{ii}(\{\mathbf{I}-\mathbf{P}\}u) + \partial_i b_i^u - U * b_i^u b_i^u = A_{ii}(l+r), \tag{2.23}$$

$$\begin{aligned}\partial_t A_{ij}(\{\mathbf{I}-\mathbf{P}\}u) + \partial_i b_j^u + \partial_j b_i^u - U * b_i^u b_j^u - U * b_j^u b_i^u \\ = A_{ij}(l+r),\ i \neq j,\end{aligned} \tag{2.24}$$

for $1 \le i, j \le n$, where l, r are defined by (2.17) and (2.18), respectively. Notice that (2.24) is symmetric in (i, j). The similar derivation of the system of equations (2.21)-(2.24) is inspired by [13] and used recently in [10] for the study of the Boltzmann equation.

The following important observation, which plays a key role in the estimates on the macroscopic dissipation firstly pointed out by [13], is that from (2.23) and (2.24), b^u satisfies the following

Proposition 2.2. *For fixed $1 \le j \le n$, it holds*

$$\begin{aligned}&\partial_t \left[\sum_{i\neq j} \partial_j A_{ii}(\{\mathbf{I}-\mathbf{P}\}u) - \sum_i \partial_i A_{ij}(\{\mathbf{I}-\mathbf{P}\}u)\right] - \Delta_x b_j^u \\ &= \sum_i \partial_j (U * b_i^u b_i^u) - \sum_i \partial_i (U * b_i^u b_j^u + U * b_j^u b_i^u) \\ &\quad + \sum_{i\neq j} \partial_j A_{ii}(l+r) - \sum_i \partial_i A_{ij}(l+r),\end{aligned} \tag{2.25}$$

for $t \ge 0$ and $x \in \mathbb{R}^n$.

Proof. For simplicity, set

$$R = -\partial_t\{\mathbf{I}-\mathbf{P}\}u + l + r.$$

From (2.24), one can compute

$$\begin{aligned}-\Delta_x b_j^u &= -\sum_{i\neq j} \partial_i(\partial_i b_j^u) - \partial_j\partial_j b_j^u \\ &= -\sum_{i\neq j} \partial_i[-\partial_j b_i^u + U * b_i^u b_j^u + U * b_j^u b_i^u + A_{ij}(R)] - \partial_j\partial_j b_j^u \\ &= \partial_j \left[\sum_{i\neq j} \partial_i b_i^u - \partial_j b_j^u\right] - \sum_{i\neq j} \partial_i[U * b_i^u b_j^u + U * b_j^u b_i^u + A_{ij}(R)].\end{aligned}$$

Thanks to (2.23), one has

$$-\Delta_x b_j^u = \sum_{i\neq j} \partial_j[U * b_i^u b_i^u + A_{ii}(R)] - \partial_j[U * b_j^u b_j^u + A_{jj}(R)]$$
$$- \sum_{i\neq j} \partial_i[U * b_i^u b_j^u + U * b_j^u b_i^u + A_{ij}(R)].$$

A further simplification gives

$$-\Delta_x b_j^u = \sum_i \partial_j(U * b_i^u b_i^u) - \sum_i \partial_i(U * b_i^u b_j^u + U * b_j^u b_i^u)$$
$$+ \sum_{i\neq j} \partial_j A_{ii}(R) - \sum_i \partial_i A_{ij}(R).$$

Then, (2.25) follows from the definition of R and the linearity of A_{ij}. This completes the proof of Proposition 2.2. □

3 Linearized Cauchy problem

3.1 Hypocoercivity

Let us now consider the Cauchy problem of the linearized equation with a nonhomogeneous source, namely

$$\begin{cases} \partial_t u = \mathbf{B}u + h, & t > 0, x \in \mathbb{R}^n, \\ u|_{t=0} = u_0, & x \in \mathbb{R}^n, \end{cases} \tag{3.1}$$

where $n \geq 1$ is the spatial dimension, $h = h(t, x, \xi)$ and $u_0 = u_0(x, \xi)$ are given, and the linear operator $\mathbf{B}$ is defined by

$$\mathbf{B} = -\xi \cdot \nabla_x + \mathbf{L}, \quad \mathbf{L} = \mathbf{L}_{FP} + \mathbf{A},$$
$$\mathbf{L}_{FP} u = \Delta_\xi u + \frac{1}{4}(2n - |\xi|^2)u,$$
$$\mathbf{A}u = U * b^u \cdot \xi\sqrt{\mathbf{M}}, \quad b^u = \langle \xi\sqrt{\mathbf{M}}, u\rangle.$$

Formally, the solution to the Cauchy problem (3.1) can be written as the Duhamel formula

$$u(t) = e^{\mathbf{B}t}u_0 + \int_0^t e^{\mathbf{B}(t-s)}h(s)ds,$$

where $e^{\mathbf{B}t}$ denotes the solution operator to the Cauchy problem of the linearized equation without source corresponding to (3.1) with $h \equiv 0$. In this section, we shall show that $e^{\mathbf{B}t}$ has the algebraic decay as time tends to infinity as in the case of the Boltzmann equation [21, 22, 12, 11].

To this end, for $1 \leq q \leq 2$ and $m \geq 0$, set the rate index $\sigma_{q,m}$ by

$$\sigma_{q,m} = \frac{n}{2}\left(\frac{1}{q} - \frac{1}{2}\right) + \frac{m}{2}.$$

The main result of this section, whose proof is left to the next subsection, is stated as follows.

Theorem 3.1. *Let $1 \leq q \leq 2$ and $n \geq 1$, and let (1.3) hold.*

(i) For any α, α' with $\alpha' \leq \alpha$, and for any u_0 satisfying $\partial_x^\alpha u_0 \in L^2_{x,\xi}$ and $\partial_x^{\alpha'} u_0 \in Z_q$, one has

$$\|\partial_x^\alpha e^{\mathbf{B}t}u_0\| \leq C(1+t)^{-\sigma_{q,m}}(\|\partial_x^{\alpha'} u_0\|_{Z_q} + \|\partial_x^\alpha u_0\|), \tag{3.2}$$

for $t \geq 0$ with $m = |\alpha - \alpha'|$, where C is a positive constant depending only on n, m, q.

(ii) Similarly, for any α, α' with $\alpha' \leq \alpha$, and for any h such that, for all $t \geq 0$ it holds $\nu(\xi)^{-1/2}\partial_x^\alpha h(t) \in L^2_{x,\xi}$, $\nu(\xi)^{-1/2}\partial_x^{\alpha'} h(t) \in Z_q$ and further

$$\int_{\mathbb{R}^n} \sqrt{\mathbf{M}} h(t,x,\xi) d\xi = \int_{\mathbb{R}^n} \xi_i \sqrt{\mathbf{M}} h(t,x,\xi) d\xi = 0, \quad i = 1,2,\ldots,n \tag{3.3}$$

$x \in \mathbb{R}^n$, one has

$$\left\| \partial_x^\alpha \int_0^t e^{\mathbf{B}(t-s)} h(s) ds \right\|^2$$
$$\leq C \int_0^t (1+t-s)^{-2\sigma_{q,m}} (\|\nu^{-1/2}\partial_x^{\alpha'} h(s)\|^2_{Z_q} + \|\nu^{-1/2}\partial_x^\alpha h(s)\|^2) ds, \tag{3.4}$$

for $t \geq 0$ with $m = |\alpha - \alpha'|$, where C is a positive constant depending only on n, m, q.

3.2 Proof of hypocoercivity: Fourier analysis

In what follows we devote ourselves to the proof of Theorem 3.1. Let $u = u(t,x,\xi)$ be the solution to the Cauchy problem (3.1) with the nonhomogeneous source $h(t,x,\xi)$ and initial data $u_0(x,\xi)$. Similarly as before, we decompose u as

$$\begin{cases} u(t,x,\xi) = \mathbf{P}u + \{\mathbf{I} - \mathbf{P}\}u, \\ \mathbf{P}u \equiv \{a^u + b^u \cdot \xi\}\sqrt{\mathbf{M}}, \\ a^u = \langle \sqrt{\mathbf{M}}, u\rangle, \quad b^u = \langle \xi\sqrt{\mathbf{M}}, u\rangle. \end{cases}$$

Then, from the same procedure as in Section 2.2, a suitable skipping of the nonlinear term $\Gamma(u,u)$ leads to the macroscopic balance laws satisfied by a^u, b^u:

$$\partial_t a^u + \nabla_x \cdot b^u = 0, \tag{3.5}$$

$$\partial_t b_i^u + \partial_i a^u - (U * b_i^u - b_i^u) + \sum_{j=1}^n \partial_j A_{ij}(\{\mathbf{I}-\mathbf{P}\}u) = 0, \tag{3.6}$$

$$\partial_t A_{ii}(\{\mathbf{I}-\mathbf{P}\}u) + \partial_i b_i^u = A_{ii}(l+h), \tag{3.7}$$

$$\partial_t A_{ij}(\{\mathbf{I}-\mathbf{P}\}u) + \partial_i b_j^u + \partial_j b_i^u = A_{ij}(l+h), \ i \neq j, \tag{3.8}$$

where $1 \leq i,j \leq n$, the velocity moment function $A_{ij}(\cdot)$ is defined by (2.20), and l has the same form as before, given by

$$l = -\xi \cdot \nabla_x \{\mathbf{I}-\mathbf{P}\}u + \mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}u. \tag{3.9}$$

Here we notice that h does not appear in the first $n+1$ equations (3.5)-(3.6) because the assumption (3.3) implies that

$$\mathbf{P}h(t,x) \equiv 0, \quad t \geq 0, x \in \mathbb{R}^n.$$

Furthermore, following a procedure similar to that used to derive (2.25) from (2.23) and (2.24), for fixed $1 \leq j \leq n$, it follows from (3.7) and (3.8) that

$$\partial_t \left[\sum_{i\neq j} \partial_j A_{ii}(\{\mathbf{I}-\mathbf{P}\}u) - \sum_i \partial_i A_{ij}(\{\mathbf{I}-\mathbf{P}\}u) \right] - \Delta_x b_j^u$$
$$= \sum_{i \neq j} \partial_j A_{ii}(l+h) - \sum_i \partial_i A_{ij}(l+h). \tag{3.10}$$

Up to the end of this subsection, let us introduce some notations. For an integrable function $g : \mathbb{R}^n \to \mathbb{R}$, its Fourier transform $\widehat{g} = \mathcal{F}g$ is defined by

$$\widehat{g}(k) = \mathcal{F}g(k) = \int_{\mathbb{R}^n} e^{-2\pi i x\cdot k} g(x) dx, \quad x \cdot k =: \sum_{j=1}^n x_j k_j.$$

Here, $k\in\mathbb{R}^n$, and $\mathrm{i}=\sqrt{-1}\in\mathbb{C}$ is the imaginary unit. For two complex vectors $a,b\in\mathbb{C}^n$, $(a\mid b)$ denotes the dot product $a\cdot\overline{b}$ over the complex field, where $\overline{b}$ is the complex conjugate of b.

Lemma 3.1. *There is a temporal-frequency free energy functional* $\mathcal{E}^l_{free}(\widehat{u}(t,k))$ *in the form of*

$$\begin{aligned}\mathcal{E}^l_{free}(\widehat{u}(t,k)) = 3\sum_j\sum_{i\neq j}&\frac{\mathrm{i}k_j}{1+|k|^2}(A_{ii}(\{\mathbf{I}-\mathbf{P}\}\widehat{u})\mid\widehat{b^u_j})\\ -3\sum_{ij}&\frac{\mathrm{i}k_i}{1+|k|^2}(A_{ij}(\{\mathbf{I}-\mathbf{P}\}\widehat{u})\mid\widehat{b^u_j})\\ -&\frac{\mathrm{i}k}{1+|k|^2}\cdot(\widehat{b^u}\mid\widehat{a^u})\end{aligned}\tag{3.11}$$

such that

$$\begin{aligned}&\frac{\partial}{\partial t}\mathrm{Re}\ \mathcal{E}^l_{free}(\widehat{u}(t,k))+\frac{|k|^2}{4(1+|k|^2)}(|\widehat{a^u}|^2+|\widehat{b^u}|^2)\\ &\le\frac{1-\mathrm{Re}\ \widehat{U}}{1+|k|^2}|\widehat{b^u}|^2+\frac{C}{1+|k|^2}\|\nu^{-1/2}\widehat{h}\|^2_{L^2_\xi}+C\|\{\mathbf{I}-\mathbf{P}\}\widehat{u}\|^2_{L^2_\xi}\end{aligned}\tag{3.12}$$

holds for $t\ge 0$ *and* $k\in\mathbb{R}^n$. *Moreover, one has the estimate*

$$|\mathcal{E}^l_{free}(\widehat{u}(t,k))|\le C\|\widehat{u}(t,k)\|^2_{L^2_\xi},\tag{3.13}$$

for $t\ge 0$ *and* $k\in\mathbb{R}^n$.

Proof. Let us first notice that

$$\mathcal{F}(\{\mathbf{I}-\mathbf{P}\}u)=\{\mathbf{I}-\mathbf{P}\}\mathcal{F}u,\quad \mathcal{F}(\mathbf{P}u)=\mathbf{P}\mathcal{F}u.$$

After taking the Fourier transform, (3.10) reads

$$\begin{aligned}&\partial_t\left[\sum_{i\neq j}\mathrm{i}k_jA_{ii}(\{\mathbf{I}-\mathbf{P}\}\widehat{u})-\sum_i\mathrm{i}k_iA_{ij}(\{\mathbf{I}-\mathbf{P}\}\widehat{u})\right]+|k|^2\widehat{b^u_j}\\ &\qquad=\sum_{i\neq j}\mathrm{i}k_jA_{ii}(\widehat{l}+\widehat{h})-\sum_i\mathrm{i}k_iA_{ij}(\widehat{l}+\widehat{h}),\end{aligned}$$

which by further taking the inner product with $\overline{\widehat{b^u_j}}$ gives

$$\begin{aligned}&\partial_t\left(\sum_{i\neq j}\mathrm{i}k_jA_{ii}(\{\mathbf{I}-\mathbf{P}\}\widehat{u})-\sum_i\mathrm{i}k_iA_{ij}(\{\mathbf{I}-\mathbf{P}\}\widehat{u})\mid\widehat{b^u_j}\right)+|k|^2|\widehat{b^u_j}|^2\\ &=\left(\sum_{i\neq j}\mathrm{i}k_jA_{ii}(\widehat{l}+\widehat{h})-\sum_i\mathrm{i}k_iA_{ij}(\widehat{l}+\widehat{h})\mid\widehat{b^u_j}\right)\\ &\quad+\left(\sum_{i\neq j}\mathrm{i}k_jA_{ii}(\{\mathbf{I}-\mathbf{P}\}\widehat{u})-\sum_i\mathrm{i}k_iA_{ij}(\{\mathbf{I}-\mathbf{P}\}\widehat{u})\mid\partial_t\widehat{b^u_j}\right).\end{aligned}\tag{3.14}$$

The first term on the r.h.s of (3.14) is estimated by

$$\begin{aligned}&\left|\left(\sum_{i\neq j}\mathrm{i}k_jA_{ii}(\widehat{l}+\widehat{h})-\sum_i\mathrm{i}k_iA_{ij}(\widehat{l}+\widehat{h})\mid\widehat{b^u_j}\right)\right|\\ &\qquad\le\frac{1}{4}|k|^2|\widehat{b^u_j}|^2+C\sum_{ij}\left(|A_{ij}(\widehat{l})|^2+|A_{ij}(\widehat{h})|^2\right).\end{aligned}\tag{3.15}$$

By (3.9), the Fourier transform $\widehat{l}$ of l is given by

$$\widehat{l} = -\mathrm{i}\xi \cdot k\{\mathbf{I}-\mathbf{P}\}\widehat{u} + \mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}\widehat{u},$$

and thus one has

$$\begin{aligned}
|A_{ij}(\widehat{l})| &= \left|\int_{\mathbb{R}^n} (\xi_i\xi_j - 1)\sqrt{\mathbf{M}}(-\mathrm{i}\xi \cdot k\{\mathbf{I}-\mathbf{P}\}\widehat{u} + \mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}\widehat{u})d\xi\right| \\
&= \left|\int_{\mathbb{R}^n} [-\mathrm{i}\xi \cdot k(\xi_i\xi_j - 1)\sqrt{\mathbf{M}} + \mathbf{L}_{FP}((\xi_i\xi_j-1)\sqrt{\mathbf{M}})]\{\mathbf{I}-\mathbf{P}\}\widehat{u}d\xi\right| \\
&\le \|-\mathrm{i}\xi \cdot k(\xi_i\xi_j - 1)\sqrt{\mathbf{M}} + \mathbf{L}_{FP}((\xi_i\xi_j-1)\sqrt{\mathbf{M}})\|_{L^2_\xi}\|\{\mathbf{I}-\mathbf{P}\}\widehat{u}\|_{L^2_\xi} \\
&\le C(1+|k|)\|\{\mathbf{I}-\mathbf{P}\}\widehat{u}\|_{L^2_\xi}.
\end{aligned} \tag{3.16}$$

Similarly one has for $\widehat{h}$

$$\begin{aligned}
|A_{ij}(\widehat{h})| &= \left|\int_{\mathbb{R}^n} (\xi_i\xi_j - 1)\sqrt{\mathbf{M}}\widehat{h}d\xi\right| \\
&\le \|\nu^{1/2}(\xi_i\xi_j - 1)\sqrt{\mathbf{M}}\|_{L^2_\xi}\|\nu^{-1/2}\widehat{h}\|_{L^2_\xi} \\
&\le C\|\nu^{-1/2}\widehat{h}\|_{L^2_\xi}.
\end{aligned} \tag{3.17}$$

Therefore, (3.15) together with (3.16) and (3.17) imply

$$\begin{aligned}
&\left|\left(\sum_{i\neq j}\mathrm{i}k_j A_{ii}(\widehat{l}+\widehat{h}) - \sum_i \mathrm{i}k_i A_{ij}(\widehat{l}+\widehat{h}) \mid \widehat{b^u_j}\right)\right| \\
&\qquad \le \frac{1}{4}|k|^2|\widehat{b^u_j}|^2 + C(1+|k|^2)\|\{\mathbf{I}-\mathbf{P}\}\widehat{u}\|^2_{L^2_\xi} + C\|\nu^{-1/2}\widehat{h}\|^2_{L^2_\xi},
\end{aligned} \tag{3.18}$$

which gives the estimate on the first term on the r.h.s. of (3.14). For the second term, one can use the Fourier transform of (3.6)

$$\partial_t\widehat{b^u_i} + \mathrm{i}k_i\widehat{a^u} - (\widehat{U}\widehat{b^u_i} - \widehat{b^u_i}) + \sum_j \mathrm{i}k_j A_{ij}(\{\mathbf{I}-\mathbf{P}\}\widehat{u}) = 0 \tag{3.19}$$

to estimate

$$\begin{aligned}
&\left(\sum_{i\neq j}\mathrm{i}k_j A_{ii}(\{\mathbf{I}-\mathbf{P}\}\widehat{u}) - \sum_i \mathrm{i}k_i A_{ij}(\{\mathbf{I}-\mathbf{P}\}\widehat{u}) \mid \partial_t\widehat{b^u_j}\right) \\
&= \Bigg(\sum_{i\neq j}\mathrm{i}k_j A_{ii}(\{\mathbf{I}-\mathbf{P}\}\widehat{u}) - \sum_i \mathrm{i}k_i A_{ij}(\{\mathbf{I}-\mathbf{P}\}\widehat{u}) \mid \\
&\qquad\qquad -\mathrm{i}k_j\widehat{a^u} + (\widehat{U}\widehat{b^u_j} - \widehat{b^u_j}) - \sum_\ell \mathrm{i}k_\ell A_{j\ell}(\{\mathbf{I}-\mathbf{P}\}\widehat{u})\Bigg) \\
&\le \delta|k|^2|\widehat{a^u}|^2 + \frac{1}{4}|k|^2|\widehat{b^u}|^2 + C_\delta(1+|k|^2)\|\{\mathbf{I}-\mathbf{P}\}\widehat{u}\|^2_{L^2_\xi},
\end{aligned} \tag{3.20}$$

where the constant $0 < \delta \le 1$ is arbitrary. In (3.20), the property $\sup_k|\widehat{U}| \le \|U\|_{L^1_x} = 1$ and

$$|A_{ij}(\{\mathbf{I}-\mathbf{P}\}\widehat{u})| \le C\|\{\mathbf{I}-\mathbf{P}\}\widehat{u}\|_{L^2_\xi}$$

have been used. From (3.14) as well as (3.18) and (3.20), one has

$$\begin{aligned}
&\partial_t\left(\sum_{i\neq j}\mathrm{i}k_j A_{ii}(\{\mathbf{I}-\mathbf{P}\}\widehat{u}) - \sum_i \mathrm{i}k_i A_{ij}(\{\mathbf{I}-\mathbf{P}\}\widehat{u}) \mid \widehat{b^u_j}\right) + \frac{1}{2}|k|^2|\widehat{b^u_j}|^2 \\
&\le \delta|k|^2|\widehat{a^u}|^2 + C_\delta(1+|k|^2)\|\{\mathbf{I}-\mathbf{P}\}\widehat{u}\|^2_{L^2_\xi} + C\|\nu^{-1/2}\widehat{h}\|^2_{L^2_\xi},
\end{aligned} \tag{3.21}$$

for $0<\delta\leq 1$ to be determined later.

To get the dissipation $|k|^2|\widehat{a^u}|^2$ as in (3.21), we take the inner product of (3.19) with $-\mathrm{i}k\overline{\widehat{a^u}}$. it holds

$$\left(-\partial_t \mathrm{i}k\cdot\widehat{b^u}\mid\widehat{a^u}\right)+|k|^2|\widehat{a^u}|^2+\left((\widehat{U}-1)\mathrm{i}k\cdot\widehat{b^u}\mid\widehat{a^u}\right)$$
$$+\left(\sum_{ij}k_ik_jA_{ij}(\{\mathbf{I}-\mathbf{P}\}\widehat{u})\mid\widehat{a^u}\right)=0. \tag{3.22}$$

Let us write

$$\left(-\partial_t \mathrm{i}k\cdot\widehat{b^u}\mid\widehat{a^u}\right)=\partial_t\left(-\mathrm{i}k\cdot\widehat{b^u}\mid\widehat{a^u}\right)+\left(\mathrm{i}k\cdot\widehat{b^u}\mid\partial_t\widehat{a^u}\right).$$

From (3.5), which implies

$$\partial_t\widehat{a^u}+\mathrm{i}k\cdot\widehat{b^u}=0,$$

one has

$$\left(\mathrm{i}k\cdot\widehat{b^u}\mid\partial_t\widehat{a^u}\right)=\left(\mathrm{i}k\cdot\widehat{b^u}\mid -\mathrm{i}k\cdot\widehat{b^u}\right)=-|k\cdot\widehat{b^u}|^2.$$

Then, the first term on the l.h.s. of (3.22) reduces to

$$\left(-\partial_t \mathrm{i}k\cdot\widehat{b^u}\mid\widehat{a^u}\right)=\partial_t\left(-\mathrm{i}k\cdot\widehat{b^u}\mid\widehat{a^u}\right)-|k\cdot\widehat{b^u}|^2.$$

Notice that

$$\mathrm{Im}\,\widehat{U}(k)=\int_{\mathbb{R}^n}U(x)\sin(k\cdot x)dx=0,$$

since U is even. Then,

$$\widehat{U}=\mathrm{Re}\,\widehat{U}$$

holds. Thus, one has the estimate on the third term on the l.h.s. of (3.22) as

$$\begin{aligned}\left|\left((\widehat{U}-1)\mathrm{i}k\cdot\widehat{b^u}\mid\widehat{a^u}\right)\right| &\leq\frac{1}{4}(1-\mathrm{Re}\,\widehat{U})|k|^2|\widehat{a^u}|^2+(1-\mathrm{Re}\,\widehat{U})|\widehat{b^u}|^2\\ &\leq\frac{1}{4}|k|^2|\widehat{a^u}|^2+(1-\mathrm{Re}\,\widehat{U})|\widehat{b^u}|^2.\end{aligned}$$

Again we used $|\mathrm{Re}\,\widehat{U}|\leq|\widehat{U}|\leq 1$. For the fourth term on the l.h.s. of (3.22), one finally has

$$\begin{aligned}\left|\left(\sum_{ij}k_ik_jA_{ij}(\{\mathbf{I}-\mathbf{P}\}\widehat{u})\mid\widehat{a^u}\right)\right| &\leq\frac{1}{4}|k|^2|\widehat{a^u}|^2+C|k|^2\sum_{ij}|A_{ij}(\{\mathbf{I}-\mathbf{P}\}\widehat{u})|^2\\ &\leq\frac{1}{4}|k|^2|\widehat{a^u}|^2+C|k|^2\|\{\mathbf{I}-\mathbf{P}\}\widehat{u}\|_{L^2_\xi}^2.\end{aligned}$$

Thus, plugging all the above estimates into (3.22) yields

$$\begin{aligned}&\partial_t\mathrm{Re}\,\left(-\mathrm{i}k\cdot\widehat{b^u}\mid\widehat{a^u}\right)+\frac{1}{2}|k|^2|\widehat{a^u}|^2\\ &\leq|k|^2|\widehat{b^u}|^2+(1-\mathrm{Re}\,\widehat{U})|\widehat{b^u}|^2+C|k|^2\|\{\mathbf{I}-\mathbf{P}\}\widehat{u}\|_{L^2_\xi}^2.\end{aligned} \tag{3.23}$$

Therefore, (3.12) follows by taking the proper linear combination of (3.21) and (3.23) with a fixed small constant $0<\delta\leq 1$ and then dividing it by $1+|k|^2$. This completes the proof of Lemma 3.1. □

Lemma 3.2. *it holds*

$$\frac{1}{2}\frac{\partial}{\partial t}\|\widehat{u}(t,k)\|_{L^2_\xi}+\lambda|\{\mathbf{I}-\mathbf{P}\}\widehat{u}|_\nu^2+(1-\mathrm{Re}\,\widehat{U})|\widehat{b^u}|^2\leq C\|\nu^{-1/2}\widehat{h}(t,k)\|_{L^2_\xi}^2, \tag{3.24}$$

for any $t\geq 0$ and $k\in\mathbb{R}^n$.

Proof. Since

$$\mathbf{L}_{FP}\mathbf{P}u = -\mathbf{P}_1 u = -b^u \cdot \xi\sqrt{\mathbf{M}},$$

we can rewrite the first equation in (3.1) as

$$\partial_t u + \xi \cdot \nabla_x u = \mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}u - (b^u - U * b^u)\cdot \xi\sqrt{\mathbf{M}} + h.$$

Taking the Fourier transform in x yields

$$\partial_t \widehat{u} + \mathrm{i}\xi \cdot k\widehat{u} = \mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}\widehat{u} - (1-\widehat{U})\widehat{b^u}\cdot \xi\sqrt{\mathbf{M}} + \widehat{h}.$$

By taking further the inner product with $\overline{\widehat{u}}$, integrating it in ξ over $\mathbb{R}^n$ and then using the coercivity estimate on $\mathbf{L}_{FP}$ (2.4), one has

$$\frac{1}{2}\frac{\partial}{\partial t}\|\widehat{u}(t)\|^2_{L^2_\xi} + \lambda|\{\mathbf{I}-\mathbf{P}\}\widehat{u}|^2_\nu + (1-\mathrm{Re}\,\widehat{U})|\widehat{b^u}|^2 \leq |\langle \widehat{h}, \overline{\widehat{u}}\rangle|, \tag{3.25}$$

where we used $\widehat{U} = \mathrm{Re}\,\widehat{U}$ and $\{\mathbf{I}-\mathbf{P}_0\}\{\mathbf{I}-\mathbf{P}\} = \{\mathbf{I}-\mathbf{P}\}$. For the r.h.s. term, since $\mathbf{P}h = 0$, it holds

$$\langle \widehat{h}, \overline{\widehat{u}}\rangle = \langle \widehat{h}, \overline{\mathbf{P}\widehat{u}}\rangle + \langle \widehat{h}, \overline{\{\mathbf{I}-\mathbf{P}\}\widehat{u}}\rangle = \langle \widehat{h}, \overline{\{\mathbf{I}-\mathbf{P}\}\widehat{u}}\rangle,$$

which implies that

$$\begin{aligned}|\langle \widehat{h}, \overline{\widehat{u}}\rangle| &\leq \delta\|\nu^{1/2}\{\mathbf{I}-\mathbf{P}\}\widehat{u}\|^2_{L^2_\xi} + \frac{1}{4\delta}\|\nu^{-1/2}\widehat{h}\|^2_{L^2_\xi}\\ &\leq C\delta|\{\mathbf{I}-\mathbf{P}\}\widehat{u}|^2_\nu + \frac{1}{4\delta}\|\nu^{-1/2}\widehat{h}\|^2_{L^2_\xi},\end{aligned} \tag{3.26}$$

where $\delta > 0$ is arbitrary. Therefore, (3.24) follows from (3.25) together with (3.26) by taking a properly small constant $\delta > 0$. This completes the proof of Lemma 3.2. □

Proof of Theorem 3.1: Let u_0 and h be given as in Theorem 3.1, and u be the solution to the Cauchy problem (3.1). Then, by choosing $M > 0$ large enough, it follows from (3.12) and (3.24) that there is an energy $\mathcal{E}^l_M(\widehat{u}(t,k))$ with

$$\mathcal{E}^l_M(\widehat{u}(t,k)) = M\|\widehat{u}(t,k)\|^2_{L^2_\xi} + \mathrm{Re}\,\mathcal{E}^l_{free}(\widehat{u}(t,k))$$

such that

$$\begin{aligned}\frac{\partial}{\partial t}\mathcal{E}^l_M(\widehat{u}(t,k)) &+ \lambda\left[|\{\mathbf{I}-\mathbf{P}\}\widehat{u}|^2_\nu + \frac{|k|^2}{1+|k|^2}(|\widehat{a^u}|^2 + |\widehat{b^u}|^2)\right]\\ &+ \lambda(1-\mathrm{Re}\,\widehat{U})|\widehat{b^u}|^2 \leq C\|\nu^{-1/2}\widehat{h}(t,k)\|^2_{L^2_\xi},\end{aligned} \tag{3.27}$$

for any $t \geq 0$ and $k \in \mathbb{R}^n$, where $\mathcal{E}^l_{free}(\widehat{u}(t,k))$ is defined by (3.11). From (3.13) it follows that

$$\mathcal{E}^l_M(\widehat{u}(t,k)) \sim \|\widehat{u}(t,k)\|^2_{L^2_\xi}, \tag{3.28}$$

if $M > 0$ is large enough. Notice that

$$|\{\mathbf{I}-\mathbf{P}\}\widehat{u}|^2_\nu + \frac{|k|^2}{1+|k|^2}(|\widehat{a^u}|^2 + |\widehat{b^u}|^2) \geq \frac{|k|^2}{1+|k|^2}\left[|\{\mathbf{I}-\mathbf{P}\}\widehat{u}|^2_\nu + |\widehat{a^u}|^2 + |\widehat{b^u}|^2\right]. \tag{3.29}$$

On the other hand, one also has

$$\begin{aligned}\mathcal{E}^l_M(\widehat{u}(t,k)) &\leq C\|\widehat{u}(t,k)\|^2_{L^2_\xi} \leq C\|\mathbf{P}\widehat{u}(t,k)\|^2_{L^2_\xi} + C\|\{\mathbf{I}-\mathbf{P}\}\widehat{u}(t,k)\|^2_{L^2_\xi}\\ &\leq C\left[|\{\mathbf{I}-\mathbf{P}\}\widehat{u}|^2_\nu + |\widehat{a^u}|^2 + |\widehat{b^u}|^2\right].\end{aligned} \tag{3.30}$$

Thus, (3.27) together with (3.29) and (3.30) yield

$$\frac{\partial}{\partial t}\mathcal{E}_M^l(\widehat{u}(t,k))+\frac{\lambda|k|^2}{1+|k|^2}\mathcal{E}_M^l(\widehat{u}(t,k))+\lambda(1-\mathrm{Re}\ \widehat{U})|\widehat{b^u}|^2\leq C\|\nu^{-1/2}\widehat{h}(t,k)\|_{L^2_\xi}^2,$$

which, by using the Gronwall inequality, gives

$$\mathcal{E}_M^l(\widehat{u}(t,k))\leq e^{-\frac{\lambda|k|^2}{1+|k|^2}t}\mathcal{E}_M^l(\widehat{u_0}(k))+C\int_0^t e^{-\frac{\lambda|k|^2}{1+|k|^2}(t-s)}\|\nu^{-1/2}\widehat{h}(s,k)\|_{L^2_\xi}^2 ds.$$

Thus, from (3.28) one obtains

$$\|\widehat{u}(t,k)\|_{L^2_\xi}^2\leq Ce^{-\frac{\lambda|k|^2}{1+|k|^2}t}\|\widehat{u_0}(k)\|_{L^2_\xi}^2+C\int_0^t e^{-\frac{\lambda|k|^2}{1+|k|^2}(t-s)}\|\nu^{-1/2}\widehat{h}(s,k)\|_{L^2_\xi}^2 ds, \tag{3.31}$$

for any $t\geq 0$ and $k\in\mathbb{R}^n$.

Now, in order to get the decay estimate (3.2), let $h=0$ so that $u(t)=e^{\mathbf{B}t}u_0$. Write $k^\alpha=k_1^{\alpha_1}k_2^{\alpha_2}\cdots k_n^{\alpha_n}$. Then, from (3.31), one has

$$\|\partial_x^\alpha e^{\mathbf{B}t}u_0\|^2=\int_{\mathbb{R}^n_k}|k^{2\alpha}|\cdot\|\widehat{u}(t,k)\|_{L^2_\xi}^2 dk\leq C\int_{\mathbb{R}^n_k}|k^{2\alpha}|e^{-\frac{\lambda|k|^2}{1+|k|^2}t}\|\widehat{u_0}(k)\|_{L^2_\xi}^2 dk. \tag{3.32}$$

As in [17], one can further estimate it by

$$\begin{aligned}
&\int_{\mathbb{R}^n_k}|k^{2\alpha}|e^{-\frac{\lambda|k|^2}{1+|k|^2}t}\|\widehat{u_0}(k)\|_{L^2_\xi}^2 dk\\
&\leq\int_{|k|\leq 1}|k^{2(\alpha-\alpha')}|e^{-\frac{\lambda|k|^2}{1+|k|^2}t}|k^{2\alpha'}|\cdot\|\widehat{u_0}(k)\|_{L^2_\xi}^2 dk+\int_{|k|\geq 1}e^{-\frac{\lambda}{2}t}|k^{2\alpha}|\cdot\|\widehat{u_0}(k)\|_{L^2_\xi}^2 dk\\
&\leq C(1+t)^{-\frac{n}{q}+\frac{n-2|\alpha-\alpha'|}{2}}\|\partial_x^{\alpha'}u_0\|_{Z_q}^2+Ce^{-\frac{\lambda}{2}t}\|\partial_x^\alpha u_0\|^2,
\end{aligned} \tag{3.33}$$

where the Hölder and Hausdorff-Young inequalities were used in the usual way. Hence, (3.2) follows from (3.32) and (3.33). On the other hand, to get the decay estimate (3.4), let $u_0=0$ so that

$$u(t)=\int_0^t e^{\mathbf{B}(t-s)}h(s)ds.$$

From (3.31), one obtains

$$\|\widehat{u}(t,k)\|_{L^2_\xi}^2\leq C\int_0^t e^{-\frac{\lambda|k|^2}{1+|k|^2}(t-s)}\|\nu^{-1/2}\widehat{h}(s,k)\|_{L^2_\xi}^2 ds. \tag{3.34}$$

Proceeding as in the derivation of (3.32) and (3.33), (3.4) follows from (3.34). This completes the proof of Theorem 3.1.

4 Nonlinear Cauchy problem

4.1 Uniform a priori estimates

From now on, we devote ourselves to the proof of the main result Theorem 1.1. Through this subsection, let u be the solution to the Cauchy problem (1.6) or equivalently (2.12) and (1.7), and let

$$n\geq 3,\quad N\geq 2[n/2]+2. \tag{4.1}$$

Also we suppose that u is smooth enough to justify that all calculations can be carried out. By using the classical energy method, we shall obtain in this subsection some uniform a priori estimates on u on the basis of some energy and energy dissipation rate inequalities. By these a priori estimates one will obtain in the next subsection a proof of the global existence of solutions with the help of the local existence as well as the continuum argument, under the smallness and regularity conditions on initial data u_0. For the time-decay rate of u, we shall apply in the last subsection the energy-spectrum method recently developed in [12] and later in [11], which combine the linearized spectral analysis given in Section 3 with the nonlinear high-order energy estimates.

For the above purpose, we begin with the proof of uniform a priori estimates on u to obtain the microscopic dissipation rate

$$\sum_{|\alpha|+|\beta|\leq N} \|\{\mathbf{I}-\mathbf{P}\}\partial_x^\alpha\partial_\xi^\beta u(t)\|_\nu^2 + \sum_{|\alpha|\leq N} \|T_\Delta \partial_x^\alpha b^u(t)\|_U^2,$$

which corresponds to the total temporal energy. Firstly, from the equation (1.6) or equivalently (2.12), one can obtain estimates on u and its space derivatives. The proof of these estimates will be postponed to Appendix A.1 for a simpler presentation. Here, we need to take care of the zero-order individually since the estimate on the nonlinear term $\Gamma(u,u)$ is a little subtle in the case of zero-order.

Lemma 4.1 (Zero-order). *it holds*

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\|u(t)\|^2 + \lambda\|\{\mathbf{I}-\mathbf{P}\}u\|_\nu^2 + \frac{1}{2}\|T_\Delta b^u\|_U^2 \\
&\leq C\|(a^u,b^u)\|_{L_x^2\cap L_x^\infty}(\|\{\mathbf{I}-\mathbf{P}\}u\|_\nu^2 + \|T_\Delta b^u\|_U^2) \\
&\quad + C\|(a^u,b^u)\|_{L_x^2}\|b^u\|_{L_x^\infty}^2,
\end{aligned} \tag{4.2}$$

for $t\geq 0$, where $\lambda>0$ and C are constants depending only on n.

Lemma 4.2 (Space derivatives). *it holds*

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\sum_{1\leq|\alpha|\leq N}\|\partial_x^\alpha u(t)\|^2 + \lambda\sum_{1\leq|\alpha|\leq N}\left(\|\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u\|_\nu^2 + \|T_\Delta\partial_x^\alpha b^u\|_U^2\right) \\
&\leq C\|\nabla_x(a^u,b^u)\|_{H_x^{N-1}}\left(\sum_{|\alpha|\leq N}\|\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u\|_\nu^2 + \|\nabla_x(a^u,b^u)\|_{H_x^{N-1}}^2\right) \\
&\quad + C\|\nabla_x b^u\|_{H_x^{N-1}}\sum_{1\leq|\alpha|\leq N-1}\|\partial_x^\alpha\nabla_\xi\{\mathbf{I}-\mathbf{P}\}u\|^2,
\end{aligned} \tag{4.3}$$

for $t\geq 0$, where $\lambda>0$ and C are constants depending only on n.

Next, we shall obtain estimates on the mixed space-velocity derivatives of u which appears on the r.h.s. of (4.3). Notice that by taking the velocity derivatives we do not affect $L_{x,\xi}^2$-norms for the macroscopic component $\mathbf{P}u$. Thus, let us apply $\mathbf{I}-\mathbf{P}$ to both sides of (2.12) to get

$$\partial_t\{\mathbf{I}-\mathbf{P}\}u + \{\mathbf{I}-\mathbf{P}\}(\xi\cdot\nabla_x u + U*b^u\cdot\nabla_\xi u) = \{\mathbf{I}-\mathbf{P}\}\mathbf{L}u + \{\mathbf{I}-\mathbf{P}\}\Gamma(u,u). \tag{4.4}$$

One can make further simplifications on the r.h.s. terms. In fact, from Theorem 2.1 (ii) it follows that

$$\begin{aligned}
\{\mathbf{I}-\mathbf{P}\}\mathbf{L}u &= \{\mathbf{I}-\mathbf{P}\}\mathbf{L}_{FP}u + \{\mathbf{I}-\mathbf{P}\}\mathbf{A}u = \{\mathbf{I}-\mathbf{P}\}\mathbf{L}_{FP}u \\
&= \mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}u + \mathbf{L}_{FP}\mathbf{P}u - \mathbf{P}\mathbf{L}_{FP}u \\
&= \mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}u - \mathbf{P}_1 u + \mathbf{P}_1 u = \mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}u.
\end{aligned}$$

Similarly, it holds

$$\begin{aligned}
\{\mathbf{I}-\mathbf{P}\}\Gamma(u,u) &= \{\mathbf{I}-\mathbf{P}\}(U*a^u\mathbf{L}_{FP}u + \frac{1}{2}U*b^u\cdot\xi u) \\
&= U*a^u\mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}u + \frac{1}{2}U*b^u\cdot\xi\{\mathbf{I}-\mathbf{P}\}u \\
&\quad + \frac{1}{2}U*b^u\cdot[\xi,\{\mathbf{I}-\mathbf{P}\}]u \\
&= \Gamma(u,\{\mathbf{I}-\mathbf{P}\}u) + \frac{1}{2}U*b^u\cdot[\xi,\{\mathbf{I}-\mathbf{P}\}]u,
\end{aligned}$$

where $\{\mathbf{I}-\mathbf{P}\}\mathbf{L}_{FP} = \mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}$ was used. Moreover $[\xi, \{\mathbf{I}-\mathbf{P}\}]$ denotes the commutator

$$[\xi_i, \{\mathbf{I}-\mathbf{P}\}] = \xi_i\{\mathbf{I}-\mathbf{P}\} - \{\mathbf{I}-\mathbf{P}\}\xi_i = [\xi_i, \mathbf{P}], \quad 1 \le i \le n,$$

with ξ regarded as the velocity multiplier operator. Therefore, (4.4) is simplified as

$$\begin{aligned}\partial_t\{\mathbf{I}-\mathbf{P}\}u + \{\mathbf{I}-\mathbf{P}\}(\xi \cdot \nabla_x u + U * b^u \cdot \nabla_\xi u) \\ = \mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}u + \Gamma(u, \{\mathbf{I}-\mathbf{P}\}u) + \frac{1}{2}U * b^u \cdot [\xi, \mathbf{P}]u,\end{aligned}$$

which further can be rewritten as the evolution equation of $\{\mathbf{I}-\mathbf{P}\}u$:

$$\begin{aligned}&\partial_t\{\mathbf{I}-\mathbf{P}\}u + \xi \cdot \nabla_x\{\mathbf{I}-\mathbf{P}\}u + U * b^u \cdot \nabla_\xi\{\mathbf{I}-\mathbf{P}\}u \\ &= \mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}u + \Gamma(u, \{\mathbf{I}-\mathbf{P}\}u) + \frac{1}{2}U * b^u \cdot [\xi, \mathbf{P}]u \\ &\quad +\mathbf{P}(\xi \cdot \nabla_x\{\mathbf{I}-\mathbf{P}\}u + U * b^u \cdot \nabla_\xi\{\mathbf{I}-\mathbf{P}\}u) \\ &\quad -\{\mathbf{I}-\mathbf{P}\}(\xi \cdot \nabla_x\mathbf{P}u + U * b^u \cdot \nabla_\xi\mathbf{P}u).\end{aligned} \tag{4.5}$$

Then, on the basis of the above equation, one can use the energy estimates to obtain the following technical lemma, which is proven in Appendix A.1.

Lemma 4.3 (Mixed space-velocity derivatives)**.** *Let $1 \le k \le N$. it holds*

$$\begin{aligned}&\frac{1}{2}\frac{d}{dt}\sum_{\substack{|\beta|=k \\ |\alpha|+|\beta|\le N}} \|\partial_x^\alpha\partial_\xi^\beta\{\mathbf{I}-\mathbf{P}\}u\|^2 + \lambda \sum_{\substack{|\beta|=k \\ |\alpha|+|\beta|\le N}} \|\partial_x^\alpha\partial_\xi^\beta\{\mathbf{I}-\mathbf{P}\}u\|_\nu^2 \\ &\le C\|(a^u, b^u)\|_{H_x^N}\left(\sum_{|\alpha|+|\beta|\le N} \|\partial_x^\alpha\partial_\xi^\beta\{\mathbf{I}-\mathbf{P}\}u\|_\nu^2 + \|\nabla_x(a^u, b^u)\|_{H_x^{N-1}}^2\right) \\ &\quad +C\left(\sum_{|\alpha|\le N-k+1} \|\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u\|_\nu^2 + \|\nabla_x(a^u, b^u)\|_{H_x^{N-k}}^2\right) \\ &\quad +C\chi_{\{2\le k\le N\}} \sum_{\substack{1\le|\beta|\le k-1 \\ |\alpha|+|\beta|\le N}} \|\partial_x^\alpha\partial_\xi^\beta\{\mathbf{I}-\mathbf{P}\}u\|_\nu^2,\end{aligned}$$

for $t \ge 0$, where $\lambda > 0$ and C are constants depending only on n, and χ_D denotes the characteristic function of a set D.

Finally, in order to control the nonlinear term and close the a priori estimates under the smallness condition, we need to obtain the macroscopic dissipation rate:

$$\sum_{|\alpha|\le N-1} \|\partial_x^\alpha\nabla_x\mathbf{P}u(t)\|^2 \sim \sum_{|\alpha|\le N-1} \|\partial_x^\alpha\nabla_x(a^u, b^u)\|^2$$

which corresponds to certain temporal free energy. Actually, the following lemma exactly gives the above dissipation for the macroscopic component $\mathbf{P}u$ or equivalently the coefficients (a^u, b^u). Here, the analysis is essentially based only on the macroscopic balance laws (2.21)-(2.24) satisfied by (a^u, b^u) which have been derived in Subsection 2.2. The proof will be carried out in the physical phase space by using a method close to the proof of Lemma 3.1 in the case of the linearized equation. Again, we postpone it to Appendix A.2.

Lemma 4.4. *There exists a temporal free energy $\mathcal{E}^n_{free}(u(t))$ of the form*

$$\begin{aligned}\mathcal{E}^n_{free}(u(t)) = 3 &\sum_{|\alpha|\leq N-1}\sum_j\sum_{i\neq j}\int_{\mathbb{R}^n} A_{ii}(\partial_x^\alpha\partial_j\{\mathbf{I}-\mathbf{P}\}u)\partial_x^\alpha b_j^u dx\\ &-3\sum_{|\alpha|\leq N-1}\sum_{ij}\int_{\mathbb{R}^n} A_{ij}(\partial_x^\alpha\partial_i\{\mathbf{I}-\mathbf{P}\}u)\partial_x^\alpha b_j^u dx\\ &+\sum_{|\alpha|\leq N-1}\int_{\mathbb{R}^n}\partial_x^\alpha\nabla_x a^u\cdot\partial_x^\alpha b^u dx,\end{aligned}\tag{4.6}$$

such that

$$\begin{aligned}&\frac{d}{dt}\mathcal{E}^n_{free}(u(t))+\lambda\|\nabla_x(a^u,b^u)\|^2_{H^{N-1}_x}\\ &\leq C\sum_{|\alpha|\leq N}(\|T_\Delta\partial_x^\alpha b^u\|^2_U+\|\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u\|^2)\\ &\quad +C\|(a^u,b^u)\|^2_{H^N_x}(\|\nabla_x(a^u,b^u)\|^2_{H^{N-1}_x}+\sum_{|\alpha|\leq N}\|\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u\|^2)\end{aligned}\tag{4.7}$$

holds for $t\geq 0$, where $\lambda>0$ and C are constants depending only on n. Moreover, it holds

$$|\mathcal{E}^n_{free}(u(t))|\leq C\|u(t)\|^2_{L^2_\xi(H^N_x)}\tag{4.8}$$

for $t\geq 0$.

We remark that an estimate similar to the one stated in Lemma 4.4 was firstly considered in [9] and recently developed in [10] in the study of the Boltzmann equation for the hard sphere model in $\mathbb{R}^n$. In addition, the proofs of Lemma 4.4 and Lemma 3.1 at the level of linearization are in the same spirit even though the analysis of the latter is made pointwise both in time and frequency.

4.2 Proof of global existence and uniqueness

In this subsection, we are going to make a few preparations in order to prove Theorem 1.1 along the line mentioned at the beginning of Subsection 4.1. Let us first consider the local existence of solutions to the Cauchy problem (1.6) or equivalently (2.12) and (1.7). We define iteratively the sequence $(f^m(t,x,\xi))_{m=0}^\infty$ of solutions to the Cauchy problems

$$\begin{cases}\partial_t f^{m+1}+\xi\cdot\nabla_x f^{m+1}+U*\rho_{\xi f^m}\cdot\nabla_\xi f^{m+1}\\ \qquad\qquad\qquad = U*\rho_{f^m}\nabla_\xi\cdot(\nabla_\xi f^{m+1}+\xi f^{m+1}),\\ f^{m+1}\equiv\mathbf{M}+\sqrt{\mathbf{M}}u^{m+1},\\ f^{m+1}|_{t=0}=f_0\equiv\mathbf{M}+\sqrt{\mathbf{M}}u_0,\end{cases}\tag{4.9}$$

or equivalently in terms of $u^m(t,x,\xi)$:

$$\begin{cases}\partial_t u^{m+1}+\xi\cdot\nabla_x u^{m+1}+U*b^{u^m}\cdot\nabla_\xi u^{m+1}\\ \qquad\qquad\qquad = \mathbf{L}_{FP}u^{m+1}+\Gamma(u^m,u^{m+1})+\mathbf{A}u^m,\\ u^{m+1}|_{t=0}=u_0,\end{cases}\tag{4.10}$$

where $m\geq 0$, and $u^0\equiv 0$ is set at the beginning of iteration. Let the solution space $X(0,T;M)$ be defined by

$$X(0,T;M)=\left\{\begin{array}{l} v\in C([0,T];H^N(\mathbb{R}^n\times\mathbb{R}^n)):\\ \sup_{0\leq t\leq T}\|v(t)\|_{H^N_{x,\xi}}\leq M,\quad \mathbf{M}+\sqrt{\mathbf{M}}v\geq 0\end{array}\right\}.$$

We prove the following

Theorem 4.1. *Let n, N satisfy (4.1). There are constants $T_*>0$, ϵ_0, M_0 such that if $u_0\in H^N(\mathbb{R}^n\times\mathbb{R}^n)$ with $f_0\equiv \mathbf{M}+\sqrt{\mathbf{M}}u_0\geq 0$ and $\|u_0\|_{H^N_{x,\xi}}\leq \epsilon_0$, then for each $m\geq 1$, u^m is well-defined with*

$$u^m\in X(0,T_*;M_0). \tag{4.11}$$

Furthermore, $(u^m)_{m\geq 0}$ is a Cauchy sequence in the Banach space $C([0,T_*];H^{N-1}(\mathbb{R}^n\times\mathbb{R}^n))$, and the corresponding limit function denoted by u belongs to $X(0,T_*;M_0)$, and u is a solution to the Cauchy problem (1.6)-(1.7). Meanwhile, there exists at most one solution in $X(0,T_*;M_0)$ to the Cauchy problem (1.6)-(1.7).

Proof. One can use induction to prove (4.11). Suppose that (4.11) holds true for $m\geq 0$. Without loss of generality, one can also suppose that u^m is smooth enough so that all the forthcoming calculations can be carried out. Otherwise, one can instead consider the Cauchy problem on the regularized iterative equation

$$\begin{aligned}&\partial_t f^{m+1,\epsilon}+\xi\cdot\nabla_x f^{m+1,\epsilon}+U*\rho_{\xi f^{m,\epsilon}}\cdot\nabla_\xi f^{m+1,\epsilon}\\&\qquad =U*\rho_{f^{m,\epsilon}}\nabla_\xi\cdot(\nabla_\xi f^{m+1,\epsilon}+\xi f^{m+1,\epsilon})+\epsilon\Delta_x f^{m+1,\epsilon},\\&f^{m+1}|_{t=0}=u_0^\epsilon.\end{aligned}$$

for any $\epsilon>0$ with u_0^ϵ a smooth approximation of u_0, prove the same for $f^{m,\epsilon}$ and then pass to the limit by letting $\epsilon\to 0$.

Thanks to the nonnegativity

$$U*\rho_{f^m}=1+U*a^{u^m},$$

one has

$$1-C_1\sqrt{M_m(T)}\leq U*\rho_{f^m}\leq 1+C_1\sqrt{M_m(T)},\quad 0\leq t\leq T,$$

for some constant $C_1>0$, where

$$M_m(T)=\sup_{0\leq t\leq T}\|u^m(t)\|^2_{H^N_{x,\xi}}$$

for any $0\leq T\leq T_*$. Note that $2C_1M_0\leq 1$ if $M_0>0$, to be chosen later, is sufficiently small. If this is the case, from the induction hypothesis, the estimate

$$2C_1\sqrt{M_m(T)}\leq 2C_1M_0\leq 1$$

holds. Then $1/2\leq U*\rho_{f^m}\leq 3/2$ follows. By the maximum principle for (4.9), one has

$$f^{m+1}\equiv\mathbf{M}+\sqrt{\mathbf{M}}u^{m+1}\geq 0.$$

To obtain the bound on u^{m+1}, for any α with $|\alpha|\leq N$, it follows from (4.10)

$$\begin{aligned}&\frac{1}{2}\frac{d}{dt}\|\partial_x^\alpha u^{m+1}(t)\|^2+\lambda_0\|\{\mathbf{I}-\mathbf{P}_0\}\partial_x^\alpha u^{m+1}\|_\nu^2\\&=\sum_{\alpha'<\alpha}C_{\alpha'}^\alpha\int_{\mathbb{R}^n}\langle U*\partial_x^{\alpha-\alpha'}b^{u^m}\cdot\nabla_\xi\partial_x^{\alpha'}u^{m+1},\partial_x^\alpha u^{m+1}\rangle dx\\&\quad+\int_{\mathbb{R}^n}\langle\mathbf{A}\partial_x^\alpha u^m,\partial_x^\alpha u^{m+1}\rangle dx+\int_{\mathbb{R}^n}\langle\partial_x^\alpha\Gamma(u^m,u^{m+1}),\partial_x^\alpha u^{m+1}\rangle dx\\&\leq C\|u^m\|_{L^2_\xi(H^N_x)}\sum_{|\alpha'|+|\beta'|\leq N}\|\partial_x^{\alpha'}\partial_\xi^{\beta'}u^{m+1}\|_\nu^2\\&\quad+C\|u^m\|_{L^2_\xi(H^N_x)}\|u^{m+1}\|_{L^2_\xi(H^N_x)}.\end{aligned}$$

Taking the summation over $|\alpha|\leq N$ implies

$$\begin{aligned}&\frac{1}{2}\frac{d}{dt}\|u^{m+1}(t)\|^2_{L^2_\xi(H^N_x)}+\frac{\lambda_0}{2}\sum_{|\alpha|\leq N}\|\partial_x^\alpha u^{m+1}\|_\nu^2\\&\leq C\|u^m\|_{L^2_\xi(H^N_x)}\sum_{|\alpha|+|\beta|\leq N}\|\partial_x^\alpha\partial_\xi^\beta u^{m+1}\|_\nu^2\\&\quad+C\|u^m\|_{L^2_\xi(H^N_x)}\|u^{m+1}\|_{L^2_\xi(H^N_x)}+C\|u^{m+1}\|^2_{L^2_\xi(H^N_x)}.\end{aligned}$$

Similarly, for any $0 \le t \le T \le T_*$, it holds

$$
\begin{aligned}
&\frac{d}{dt}\|u^{m+1}(t)\|^2_{H^N_{x,\xi}} + \lambda_0 \sum_{|\alpha|+|\beta|\le N} \|\partial^\alpha_x \partial^\beta_\xi u^{m+1}\|^2_\nu \\
&\le C\|u^m\|_{L^2_\xi(H^N_x)} \sum_{|\alpha|+|\beta|\le N} \|\partial^\alpha_x \partial^\beta_\xi u^{m+1}\|^2_\nu \\
&\quad + C\|u^m\|_{H^N_{x,\xi}} \|u^{m+1}\|_{H^N_{x,\xi}} + C\|u^{m+1}\|^2_{L^2_\xi(H^N_x)} \\
&\le C_2\sqrt{M_m(T)} \sum_{|\alpha|+|\beta|\le N} \|\partial^\alpha_x \partial^\beta_\xi u^{m+1}\|^2_\nu + C_3 M_m(T) + C_4 M_{m+1}(T),
\end{aligned} \tag{4.12}
$$

for some constant $C_2 > 0$. By letting $C_2\sqrt{M_m(T)} \le C_2 M_0 \le \lambda_0/2$, from the induction hypothesis and taking time integration, the above inequality gives

$$
\begin{aligned}
M_{m+1}(T) &+ \frac{\lambda_0}{2} \sum_{|\alpha|+|\beta|\le N} \int_0^T \|\partial^\alpha_x \partial^\beta_\xi u^{m+1}(s)\|^2_\nu ds \\
&\le \|u_0\|^2_{H^N_{x,\xi}} + C_3 M_m(T) T + C_4 M_{m+1}(T) T.
\end{aligned} \tag{4.13}
$$

Now, one can choose

$$
M_0 = \min\{\frac{1}{2C_1}, \frac{\lambda_0}{2C_2}\}, \quad T_* = \min\{\frac{1}{4C_3}, \frac{1}{2C_4}\}, \quad \epsilon_0 = \frac{1}{2}M_0
$$

so that

$$
M_{m+1}(T_*) \le 2\epsilon_0^2 + 2C_3 T_* M_m(T_*) \le \frac{1}{2}M_0^2 + \frac{1}{2}M_0^2 \le M_0^2,
$$

that is,

$$
\sup_{0\le t\le T_*} \|u^{m+1}(t)\|^2_{H^N_{x,\xi}} = \sqrt{M_{m+1}(T_*)} \le M_0.
$$

Finally, proceeding as in the proof of (4.12), for any $0 \le s \le t \le T_*$, we obtain

$$
\begin{aligned}
&\left| \|u^{m+1}(t)\|^2_{H^N_{x,\xi}} - \|u^{m+1}(s)\|^2_{H^N_{x,\xi}} \right| = \left| \int_s^t \frac{d}{d\theta} \|u^{m+1}(\theta)\|^2_{H^N_{x,\xi}} d\theta \right| \\
&\le C(M_0+1) \sum_{|\alpha|+|\beta|\le N} \int_s^t \|\partial^\alpha_x \partial^\beta_\xi u^{m+1}(\theta)\|^2_\nu d\theta + CM_0^2|t-s|.
\end{aligned} \tag{4.14}
$$

This implies that $\|u^{m+1}(t)\|^2_{H^N_{x,\xi}}$ is continuous over $0 \le t \le T_*$ since from (4.13), $\|\partial^\alpha_x \partial^\beta_\xi u^{m+1}\|^2_\nu$ is integrable over $[0, T_*]$. Hence equation (4.11) holds true for $m+1$ and so it does for any $m \ge 0$.

Next, the difference between two subsequent solutions of (4.10) satisfies

$$
\begin{cases}
\partial_t(u^{m+1}-u^m) + \xi\cdot\nabla_x(u^{m+1}-u^m) + U * b^{u^m}\cdot\nabla_\xi(u^{m+1}-u^m) \\
\qquad = \mathbf{L}_{FP}(u^{m+1}-u^m) + \Gamma(u^m, u^{m+1}-u^m) + \Gamma(u^m-u^{m-1}, u^m) \\
\qquad\quad + \mathbf{A}(u^m-u^{m-1}) - U * b^{u^m-u^{m-1}}\cdot\nabla_\xi u^m, \\
(u^{m+1}-u^m)|_{t=0} = 0.
\end{cases}
$$

As for (4.12), it follows that

$$
\begin{aligned}
&\frac{d}{dt}\|u^{m+1}(t)-u^m(t)\|^2_{H^{N-1}_{x,\xi}}+\lambda_0\sum_{|\alpha|+|\beta|\leq N-1}\|\partial_x^\alpha\partial_\xi^\beta(u^{m+1}-u^m)\|_\nu^2\\
&\leq C\|u^m\|_{L^2_\xi(H^{N-1}_x)}\sum_{|\alpha|+|\beta|\leq N-1}\|\partial_x^\alpha\partial_\xi^\beta(u^{m+1}-u^m)\|_\nu^2\\
&\quad +C\|u^m-u^{m-1}\|_{L^2_\xi(H^{N-1}_x)}\sum_{|\alpha|+|\beta|\leq N}\|\partial_x^\alpha\partial_\xi^\beta u^m\|_\nu\sum_{|\alpha|+|\beta|\leq N-1}\|\partial_x^\alpha\partial_\xi^\beta(u^{m+1}-u^m)\|_\nu\\
&\quad +C\|u^m-u^{m-1}\|_{L^2_\xi(H^{N-1}_x)}\|u^{m+1}-u^m\|_{L^2_\xi(H^{N-1}_x)}\\
&\quad +C\|u^{m+1}-u^m\|^2_{L^2_\xi(H^{N-1}_x)},
\end{aligned}
$$

where $N\geq 2[n/2]+2$ and the Sobolev embedding $H^{[n/2]+1}(\mathbb{R}^n)\hookrightarrow L^\infty(\mathbb{R}^n)$ were used. Since $\|u_0\|_{H^N_{x,\xi}}$ and hence ϵ_0, T_*, M_0 can be small enough, and from (4.13),

$$
\sup_m\int_0^{T_*}\sum_{|\alpha|+|\beta|\leq N}\|\partial_x^\alpha\partial_\xi^\beta u^m(s)\|_\nu^2 ds
$$

can be also small enough, it further follows that there is a constant $\mu<1$ such that

$$
\sup_{0\leq t\leq T_*}\|u^{m+1}(t)-u^m(t)\|_{H^{N-1}_{x,\xi}}\leq \mu\sup_{0\leq t\leq T_*}\|u^m(t)-u^{m-1}(t)\|_{H^{N-1}_{x,\xi}}. \tag{4.15}
$$

It can be seen from (4.15) that $(u^m)_{m\geq 0}$ is a Cauchy sequence in the Banach space $C([0,T_*];H^{N-1}(\mathbb{R}^n\times\mathbb{R}^n))$, and thus the limit function

$$
u\in C([0,T_*];H^{N-1}(\mathbb{R}^n\times\mathbb{R}^n))
$$

exists. By letting $m\to\infty$ in (4.9) or (4.10), u is a solution to the Cauchy problem (1.6)-(1.7). From the pointwise convergence of u^m to u by the Sobolev embedding theorem and the lower semicontinuity of the norms, $u^m\in X(0,T_*;M_0)$ implies

$$
f\equiv \mathbf{M}+\sqrt{\mathbf{M}}u\geq 0,\quad \sup_{0\leq t\leq T_*}\|u(t)\|_{H^N_{x,\xi}}\leq M_0.
$$

Similarly to the proof of (4.14), one can conclude that $u\in C([0,T_*];H^N(\mathbb{R}^n\times\mathbb{R}^n))$. Thus, $u\in X(0,T_*;M_0)$ follows.

Finally, let $v\in X(0,T_*;M_0)$ be another solution to the Cauchy problem (1.6)-(1.7). Proceeding as in the proof of (4.15) we obtain

$$
\sup_{0\leq t\leq T_*}\|u(t)-v(t)\|\leq\mu\sup_{0\leq t\leq T_*}\|u(t)-v(t)\|,
$$

for $\mu<1$. Then, $u\equiv v$, and uniqueness follows. This completes the proof of Theorem 4.1. □

Proof of global existence and uniqueness in Theorem 1.1: At this time it suffices to obtain the uniform a priori estimates. For a given $T>0$, let u be the solution to the Cauchy problem (1.6)-(1.7) over $[0,T]$ which satisfies

$$
\sup_{0\leq t\leq T}\|u(t)\|_{H^N_{x,\xi}}\leq\epsilon
$$

for $0<\epsilon\leq 1$ small enough. Now, one can apply Lemmas 4.1, 4.2, 4.3 and 4.4 to u.

We claim that there are the equivalent energy $\mathcal{E}(u(t))$ and energy dissipation rate $\mathcal{D}(u(t))$, defined by

$$
\mathcal{E}(u(t))\sim\|u(t)\|^2_{H^N_{x,\xi}}, \tag{4.16}
$$

$$
\begin{aligned}
\mathcal{D}(u(t))=&\sum_{|\alpha|+|\beta|\leq N}\|\partial_x^\alpha\partial_\xi^\beta\{\mathbf{I}-\mathbf{P}\}u(t)\|_\nu^2+\sum_{|\alpha|\leq N}\|T_\Delta\partial_x^\alpha b^u(t)\|_U^2\\
&+\|\nabla_x(a^u,b^u)(t)\|^2_{H^{N-1}_x},
\end{aligned}\tag{4.17}
$$

such that

$$\frac{d}{dt}\mathcal{E}(u(t)) + \lambda\mathcal{D}(u(t)) \leq 0, \tag{4.18}$$

holds for any $0 \leq t \leq T$. In fact, since $0 < \epsilon \leq 1$ is small enough, the linear combination of (4.2), (4.3) and (4.7) gives the dissipation of the macroscopic component $\mathbf{P}u$ or equivalently of its coefficients (a^u, b^u), the microscopic component $\{\mathbf{I}-\mathbf{P}\}u$, and their space derivatives with remaining terms including L^2-norms of the mixed space-velocity derivatives with small-coefficients, that is,

$$\begin{aligned}
&\frac{d}{dt}\left(M\|u(t)\|^2_{L^2_\xi(H^N_x)} + \mathcal{E}^n_{free}(u(t))\right)\\
&\qquad +\lambda\sum_{|\alpha|\leq N}\|\partial^\alpha_x\{\mathbf{I}-\mathbf{P}\}u\|^2_\nu + \lambda\|T_\Delta\partial^\alpha_x b^u\|^2_U + \lambda\|\nabla_x(a^u,b^u)\|^2_{H^{N-1}_x}\\
&\leq C\epsilon\mathcal{D}(u(t)) + C\epsilon\sum_{1\leq|\alpha|\leq N-1}\|\partial^\alpha_x\nabla_\xi\{\mathbf{I}-\mathbf{P}\}u\|^2,
\end{aligned} \tag{4.19}$$

where $\mathcal{E}^n_{free}(u(t))$ is defined by (4.6), and $M \geq 1$ is large enough, so that, by (4.8),

$$M\|u(t)\|^2_{L^2_\xi(H^N_x)} + \mathcal{E}^n_{free}(u(t)) \sim \|u(t)\|^2_{L^2_\xi(H^N_x)}$$

holds. On the other hand, the linear combination of (4.6) over $1 \leq k \leq N$ gives the dissipation of all the space-velocity derivatives in L^2-norm, that is,

$$\begin{aligned}
&\frac{d}{dt}\sum_{1\leq k\leq N} C_k \sum_{\substack{|\beta|=k\\|\alpha|+|\beta|\leq N}}\|\partial^\alpha_x\partial^\beta_\xi\{\mathbf{I}-\mathbf{P}\}u\|^2 + \lambda\sum_{\substack{|\beta|\geq 1\\|\alpha|+|\beta|\leq N}}\|\partial^\alpha_x\partial^\beta_\xi\{\mathbf{I}-\mathbf{P}\}u\|^2_\nu\\
&\leq C\epsilon\mathcal{D}(u(t)) + C\sum_{|\alpha|\leq N}\|\partial^\alpha_x\{\mathbf{I}-\mathbf{P}\}u\|^2_\nu + C\|\nabla_x(a^u,b^u)\|^2_{H^{N-1}_x},
\end{aligned} \tag{4.20}$$

for some properly chosen constants C_k. Thus, the further linear combination of (4.19) and (4.20) leads to (4.18) by letting $\epsilon > 0$ small enough. Then, after taking time integration, it holds

$$\sup_{0\leq t\leq T}\left\{\|u(t)\|_{H^N_{x,\xi}} + \lambda\int_0^t \mathcal{D}(u(s))ds\right\} \leq \mathcal{E}(u_0) \leq C\|u_0\|^2_{H^N_{x,\xi}},$$

where C is independent of T and u_0. Thus, the global existence and uniqueness of solutions to the Cauchy problem (1.6)-(1.7) follows from the above uniform a priori estimate (4.20) together with the local existence obtained in Theorem 4.1 as well as the continuum argument, and moreover, (1.13) and (1.14) hold. Here, the details are omitted for simplicity. This completes the proof of global existence and uniqueness in Theorem 1.1.

4.3 Proof of rates of convergence

In this subsection, in order to prove (1.15) in Theorem 1.1, we are concerned with the time-decay rates of solutions. The main idea of the proof is based on the energy-spectrum method recently developed in [12, 11]. To this end, let us suppose that all conditions in Theorem 1.1 hold, and let u be the solution to the Cauchy problem (1.6)-(1.7) satisfying (1.13) and (1.14).

Firstly, the time-decay properties of the linearized solution operator $e^{\mathbf{B}t}$ given in Section 3 will be applied in the following lemma to obtain some formal time-decay estimates for the solution u in terms of the total temporal energy $\mathcal{E}(u(t))$.

Lemma 4.5. *If $\|u_0\|_{Z_1}$ is bounded, then it holds*

$$\begin{aligned}
\|u(t)\|^2 &\leq C(\mathcal{E}(u_0) + \|u_0\|^2_{Z_1})(1+t)^{-\frac{n}{2}}\\
&\quad +C\int_0^t (1+t-s)^{-\frac{n}{2}}\mathcal{E}(u(s))[\mathcal{E}(u(s)) + \|\xi\{\mathbf{I}-\mathbf{P}\}u(s)\|^2]ds\\
&\quad +C\left[\int_0^t (1+t-s)^{-\frac{n}{4}}\mathcal{E}(u(s))ds\right]^2,
\end{aligned} \tag{4.21}$$

for any $t \geq 0$.

Proof. From (2.12), u can be written in mild form as

$$u(t) = e^{\mathbf{B}t}u_0 + \int_0^t e^{\mathbf{B}(t-s)}G(s)ds, \tag{4.22}$$

where the source term G is denoted by

$$G = \Gamma(u,u) - U * b^u \cdot \nabla_\xi u.$$

By the definition of Γ, given in (2.13), G can be rewritten as

$$G = G_1 + G_2 + G_3,$$

with

$$\begin{aligned}
G_1 &= U * a^u \mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}u,\\
G_2 &= \frac{1}{2}U * b^u \cdot \xi\{\mathbf{I}-\mathbf{P}\}u - U * b^u \cdot \nabla_\xi\{\mathbf{I}-\mathbf{P}\}u,\\
G_3 &= -U * a^u \mathbf{P}_1 u + \frac{1}{2}U * b^u \cdot \xi\mathbf{P}u - U * b^u \cdot \nabla_\xi \mathbf{P}u.
\end{aligned}$$

It is straightforward to check that both G_1 and G_2 satisfy condition (3.3). Then, one can apply both (i) and (ii) in Theorem 3.1 to (4.22) to obtain

$$\begin{aligned}
\|u(t)\|^2 \leq{}& C(\mathcal{E}(u_0) + \|u_0\|_{Z_1}^2)(1+t)^{-\frac{n}{2}}\\
&+C\sum_{i=1}^2 \int_0^t (1+t-s)^{-\frac{n}{2}}(\|\nu^{-1/2}G_i(s)\|_{Z_1}^2 + \|\nu^{-1/2}G_i(s)\|^2)ds\\
&+C\left[\int_0^t (1+t-s)^{-\frac{n}{4}}(\|G_3(s)\|_{Z_1} + \|G_3(s)\|)ds\right]^2.
\end{aligned} \tag{4.23}$$

By using the inequalities

$$\begin{aligned}
&\|U * (a^u,b^u)\|_{L^2_x} \leq \|U\|_{L^1_x}\|(a^u,b^u)\|_{L^2_x},\\
&\|U * (a^u,b^u)\|_{L^\infty_x} \leq \|U\|_{L^1_x}\|(a^u,b^u)\|_{L^\infty_x} \leq C\|U\|_{L^1_x}\|\nabla_x(a^u,b^u)\|_{H_x^{[n/2]}},
\end{aligned}$$

it is straightforward to get

$$\begin{aligned}
&\|\nu^{-1/2}G_1\|_{Z_1}^2 + \|\nu^{-1/2}G_1\|^2 \leq C\mathcal{E}(u)[\mathcal{E}(u) + \|\xi\{\mathbf{I}-\mathbf{P}\}u\|^2],\\
&\|\nu^{-1/2}G_2\|_{Z_1}^2 + \|\nu^{-1/2}G_2\|^2 \leq C[\mathcal{E}(u)]^2,\\
&\|G_3(s)\|_{Z_1} + \|G_3(s)\| \leq \mathcal{E}(u).
\end{aligned}$$

Plugging the above estimates into (4.23) leads to (4.21). This completes the proof of Lemma 4.5. □

The next lemma is devoted to obtain a uniform bound on the velocity-weighted norm $\|\xi\{\mathbf{I}-\mathbf{P}\}u(s)\|$, under the additional condition on initial data which imply that the time integral term on the r.h.s. of (4.21) can be controlled. This, together with (4.18) can lead to the desired time-decay rates of the total temporal energy $\mathcal{E}(u(t))$.

Lemma 4.6. *If* $\|u_0\|_{H^N_{x,\xi}}$ *is small enough and* $\|\xi u_0\|$ *is bounded, then it holds*

$$\|\{\mathbf{I}-\mathbf{P}\}u\|_\nu \leq C(\|u_0\|_{H^N_{x,\xi}} + \|\xi u_0\|), \tag{4.24}$$

for any $t \geq 0$.

Proof. For simplicity, set $w_i = \xi_i\{\mathbf{I}-\mathbf{P}\}u$. Thanks to (4.5), w_i satisfies

$$\begin{aligned}
&\partial_t w_i + \xi\cdot\nabla_x w_i + U * b^u \cdot \nabla_\xi w_i \\
&= \mathbf{L}_{FP} w_i + \Gamma(u, w_i) + \frac{1}{2} U * b^u \cdot \xi_i[\xi, \mathbf{P}]u \\
&\quad + \xi\mathbf{P}(\xi\cdot\nabla_x\{\mathbf{I}-\mathbf{P}\}u + U * b^u\cdot\nabla_\xi\{\mathbf{I}-\mathbf{P}\}u) \\
&\quad - \xi\{\mathbf{I}-\mathbf{P}\}(\xi\cdot\nabla_x\mathbf{P}u + U * b^u\cdot\nabla_\xi\mathbf{P}u) \\
&\quad + U * b_i^u\{\mathbf{I}-\mathbf{P}\}u - 2(1+U*a^u)\partial_{\xi_i}\{\mathbf{I}-\mathbf{P}\}u.
\end{aligned}$$

The last line follows from the computation of commutators

$$[\nabla_\xi, \xi_i] = e_i, \quad [\mathbf{L}_{FP}, \xi_i] = [\Delta_\xi, \xi_i] = 2\partial_{\xi_i}.$$

The zero-order energy estimate as before gives

$$\frac{1}{2}\frac{d}{dt}\|w_i(t)\|^2 + \lambda_0\|\{\mathbf{I}-\mathbf{P}_0\}w_i\|_\nu^2 \le \left(\frac{\lambda_0}{8} + C\mathcal{E}(u)\right)\|w_i\|_\nu^2 + C(\mathcal{E}(u)+1)\mathcal{D}(u),$$

where $\mathcal{E}(u)$, $\mathcal{D}(u)$ are defined in (4.16) and (4.17), respectively. From (4.18) and

$$\sup_{t\ge 0}\mathcal{E}(u(t)) \le \mathcal{E}(u_0) \le C\|u_0\|_{H^N_{x,\xi}}^2,$$

which is small enough, it follows

$$\frac{d}{dt}\|w_i(t)\|^2 + \lambda_0\|w_i\|_\nu^2 \le C\mathcal{D}(u) + C\|\mathbf{P}_0 w_i\|_\nu^2 \le C\mathcal{D}(u).$$

Then, further taking time integration and using (4.18), one has

$$\begin{aligned}
\|w_i(t)\|^2 + \lambda_0\int_0^t \|w_i(s)\|_\nu^2 ds &\le \|w_i(0)\|^2 + C\int_0^t \mathcal{D}(u(s))ds \\
&\le C\|\xi_i\{\mathbf{I}-\mathbf{P}\}u_0\|^2 + C\mathcal{E}(u_0) \\
&\le C(\|u_0\|_{H^N_{x,\xi}}^2 + \|\xi u_0\|^2),
\end{aligned}$$

which gives (4.24). This completes the proof of Lemma 4.6. □

Proof of time-decay rates in Theorem 1.1: For simplicity, let us denote

$$K_0 = \|u_0\|_{H^N_{x,\xi}} + \|u_0\|_{Z_1}, \quad \delta_0 = \|u_0\|_{H^N_{x,\xi}} + \|\xi u_0\|. \tag{4.25}$$

Notice that K_0 is finite and δ_0 can be arbitrarily small thanks to the assumptions of Theorem 1.1. In order to get the time decay of the total temporal energy in (1.15), we define

$$\mathcal{E}_\infty(t) = \sup_{0\le s\le t}(1+s)^{\frac{n}{2}}\mathcal{E}(u(s)).$$

Thus, to prove (1.15), it suffices to prove that $\mathcal{E}_\infty(t)$ is uniformly bounded in time. In fact, combining (4.21) and (4.24) gives

$$\begin{aligned}
\|u(t)\|^2 &\le CK_0^2(1+t)^{-\frac{n}{2}} + C\delta_0^2\int_0^t(1+t-s)^{-\frac{n}{2}}\mathcal{E}(u(s))ds \\
&\quad + C\delta_0^{\frac{2}{3}-2\epsilon}\left(\int_0^t(1+t-s)^{-\frac{n}{4}}[\mathcal{E}(u(s))]^{\frac{2}{3}+\epsilon}ds\right)^2 \\
&\le CK_0^2(1+t)^{-\frac{n}{2}} + C\delta_0^2\mathcal{E}_\infty(t)\int_0^t(1+t-s)^{-\frac{n}{2}}(1+s)^{-\frac{n}{2}}ds \\
&\quad + C\delta_0^{\frac{2}{3}-2\epsilon}[\mathcal{E}_\infty(t)]^{\frac{2}{3}+\epsilon}\left(\int_0^t(1+t-s)^{-\frac{n}{4}}(1+s)^{-\frac{n}{3}-\frac{n}{2}\epsilon}ds\right)^2,
\end{aligned} \tag{4.26}$$

where $0 < \epsilon < 1/3$ is a constant. Since $n \geq 3$ holds, one has

$$\int_0^t (1+t-s)^{-\frac{n}{2}}(1+s)^{-\frac{n}{2}} ds \leq C(1+t)^{-\frac{n}{2}},$$
$$\int_0^t (1+t-s)^{-\frac{n}{4}}(1+s)^{-\frac{n}{3}-\frac{n}{2}\epsilon} ds \leq C(1+t)^{-\frac{n}{4}},$$

where $n/2 > 1$ and $n/3 + n\epsilon/2 > \max\{1, n/4\}$ were used. Then, it follows from (4.26) that

$$\|u(t)\|^2 \leq C\left\{K_0^2 + \delta_0^2\mathcal{E}_\infty(t) + C\delta_0^{\frac{2}{3}-2\epsilon}[\mathcal{E}_\infty(t)]^{\frac{2}{3}+\epsilon}\right\}(1+t)^{-\frac{n}{2}}. \tag{4.27}$$

Notice that (4.18) implies

$$\frac{d}{dt}\mathcal{E}(u(t)) + \lambda\mathcal{E}(u(t)) \leq C\|u(t)\|^2. \tag{4.28}$$

By the Gronwall inequality, it follows from (4.27) and (4.28) that

$$\begin{aligned}\mathcal{E}(u(t)) &\leq e^{-\lambda t}\mathcal{E}(u_0) + C\int_0^t e^{-\lambda(t-s)}\|u(s)\|^2 ds \\ &\leq C\left\{K_0^2 + \delta_0^2\mathcal{E}_\infty(t) + C\delta_0^{\frac{2}{3}-2\epsilon}[\mathcal{E}_\infty(t)]^{\frac{2}{3}+\epsilon}\right\}(1+t)^{-\frac{n}{2}},\end{aligned}$$

for any $t \geq 0$. In fact $\mathcal{E}_\infty(t)$ is nondecreasing in time. Then, it holds

$$\mathcal{E}_\infty(t) \leq CK_0^2 + C\delta_0^2\mathcal{E}_\infty(t) + C\delta_0^{\frac{2}{3}-2\epsilon}[\mathcal{E}_\infty(t)]^{\frac{2}{3}+\epsilon},$$

for any $t \geq 0$. Since δ_0 in (4.25) is small enough, one has

$$\mathcal{E}_\infty(t) \leq CK_0^2 + C\delta_0^{\frac{2}{3}-2\epsilon}[\mathcal{E}_\infty(t)]^{\frac{2}{3}+\epsilon},$$

for any $t \geq 0$. Again using the smallness of δ_0 and $2/3 - 2\epsilon > 0$, one further has

$$\sup_{t\geq 0}\mathcal{E}_\infty(t) \leq CK_0^2.$$

Thus, the uniform boundness of $\mathcal{E}_\infty(t)$ is obtained and hence (1.15) is proved. This completes the proof of time decay rates in Theorem 1.1 and thus the proof of the whole Theorem 1.1.

A Proofs of uniform a priori estimates

A.1 A priori estimates: Microscopic dissipation

In the first part of this appendix, we shall prove Lemmas 4.1, 4.2 and 4.3 which are related to the microscopic dissipation rate. As a preparation, we first obtain a lemma about some estimate on the nonlinear term.

Lemma A.1. *it holds*

$$\begin{aligned}&|\langle\Gamma(u,v),w\rangle| \\ &\leq C|U*(a^u,b^u)|\left(|\{\mathbf{I}-\mathbf{P}\}v|_\nu + |(a^v,b^v)|\right)\left(|\{\mathbf{I}-\mathbf{P}\}w|_\nu + |(a^w,b^w)|\right),\end{aligned} \tag{A.1}$$

for some constant C depending only on n.

Proof. Recall the definition (2.13) of $\Gamma(u,v)$. One has

$$\begin{aligned}\langle\Gamma(u,v),w\rangle &= U*a^u\left(\langle\mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}v,\{\mathbf{I}-\mathbf{P}\}w\rangle - \langle\mathbf{P}_1 v,\mathbf{P}_1 w\rangle\right) \\ &\quad + \frac{1}{2}U*b^u\cdot\langle\xi,(\{\mathbf{I}-\mathbf{P}\}v+\mathbf{P}v)(\{\mathbf{I}-\mathbf{P}\}w+\mathbf{P}w)\rangle.\end{aligned} \tag{A.2}$$

Then, (A.1) follows by applying integration by parts and Hölder inequality to (A.1). □

Proof of Lemma 4.1: From (2.12), the zero-order energy estimate over $\mathbb{R}^n_x \times \mathbb{R}^n_\xi$ gives

$$\frac{1}{2}\frac{d}{dt}\|u(t)\|^2 - \int_{\mathbb{R}^n} \langle \mathbf{L}u, u\rangle dx = \int_{\mathbb{R}^3} \langle \Gamma(u,u), u\rangle dx. \tag{A.3}$$

To estimate the nonlinear term on the r.h.s., let us define

$$I(u; v, w) = \langle \Gamma(u, v), w\rangle,$$

where $\Gamma(\cdot,\cdot)$ means the bilinear operator given in (2.13). Notice that $I(u; v, w)$ can be written as the summation of two terms:

$$I(u; v, w) = I_1(u; v, w) + I_2(u; v, w),$$

where as in (A.2), from the macro-micro decomposition (2.10), I_1, I_2 are defined by

$$\begin{aligned} I_1(u; v, w) &= U * a^u \langle \mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}v, \{\mathbf{I}-\mathbf{P}\}w\rangle \\ &\quad + \frac{1}{2} U * b^u \cdot \langle \xi, \{\mathbf{I}-\mathbf{P}\}v\{\mathbf{I}-\mathbf{P}\}w \\ &\qquad\qquad + \{\mathbf{I}-\mathbf{P}\}v\mathbf{P}w + \mathbf{P}v\{\mathbf{I}-\mathbf{P}\}w\rangle, \\ I_2(u; v, w) &= -U * a^u \langle \mathbf{P}_1 v, \mathbf{P}_1 w\rangle + \frac{1}{2} U * b^u \cdot \langle \xi, \mathbf{P}v\mathbf{P}w\rangle. \end{aligned}$$

One can further simplify the form of $I_2(u; v, w)$ in terms of the coefficients of $\mathbf{P}v$ and $\mathbf{P}w$. In fact, it holds

$$\langle \mathbf{P}_1 v, \mathbf{P}_1 w\rangle = \int_{\mathbb{R}^n} b^v \cdot \xi b^w \cdot \xi \mathbf{M} d\xi = b^v \cdot b^w,$$

and

$$\langle \xi, \mathbf{P}v\mathbf{P}w\rangle = \int_{\mathbb{R}^n} \xi (a^v + b^v \cdot \sqrt{\mathbf{M}})(a^w + b^w \cdot \sqrt{\mathbf{M}}) d\xi = a^v b^w + a^w b^v.$$

Then, it follows that

$$I_2(u; v, w) = -U * a^u b^v \cdot b^w + \frac{1}{2} U * b^u \cdot (a^v b^w + a^w b^v).$$

Next, we estimate the nonlinear term

$$\langle \Gamma(u,u), u\rangle = I_1(u; u, u) + I_2(u; u, u). \tag{A.4}$$

For $I_1(u; u, u)$, one has

$$\begin{aligned} \int_{\mathbb{R}^n} I_1(u; u, u) dx &= \int_{\mathbb{R}^n} U * a^u \langle \mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}u, \{\mathbf{I}-\mathbf{P}\}u\rangle dx \\ &\quad + \int_{\mathbb{R}^n} U * b^u \cdot \langle \xi, \mathbf{P}u\{\mathbf{I}-\mathbf{P}\}u\rangle dx \\ &\quad + \frac{1}{2}\int_{\mathbb{R}^n} U * b^u \cdot \langle \xi, |\{\mathbf{I}-\mathbf{P}\}u|^2\rangle dx. \end{aligned}$$

The terms on the r.h.s are estimated as

$$\begin{aligned} &\int_{\mathbb{R}^n} U * a^u \langle \mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}u, \{\mathbf{I}-\mathbf{P}\}u\rangle dx \\ &\le \|U * a^u\|_{L^\infty_x} \int_{\mathbb{R}^n} |\langle \mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}u, \{\mathbf{I}-\mathbf{P}\}u\rangle| dx \\ &\le C\|a^u\|_{L^\infty_x} \|\{\mathbf{I}-\mathbf{P}\}u\|_\nu^2, \end{aligned}$$

$$
\begin{aligned}
&\int_{\mathbb{R}^n} U * b^u \cdot \langle \xi, \mathbf{P}u\{\mathbf{I}-\mathbf{P}\}u\rangle dx \\
&\leq \int_{\mathbb{R}^n} |U * b^u| \cdot \|\xi \mathbf{P}u\|_{L^2_\xi} \cdot \|\{\mathbf{I}-\mathbf{P}\}u\|_{L^2_\xi} dx \\
&\leq C\int_{\mathbb{R}^n} |U * b^u| \cdot (|a^u| + |b^u|) \cdot \|\{\mathbf{I}-\mathbf{P}\}u\|_{L^2_\xi} dx \\
&\leq C\|(a^u, b^u)\|_{L^2_x} \|b^u\|_{L^\infty_x} \|\{\mathbf{I}-\mathbf{P}\}u\|_\nu,
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{1}{2}\int_{\mathbb{R}^n} U * b^u \cdot \langle \xi, |\{\mathbf{I}-\mathbf{P}\}u|^2\rangle dx &\leq \frac{1}{2}\int_{\mathbb{R}^n} |U * b^u| \cdot |\{\mathbf{I}-\mathbf{P}\}u|^2_\xi dx \\
&\leq \frac{1}{2}\|b^u\|_{L^\infty_x} \|\{\mathbf{I}-\mathbf{P}\}u\|^2_\nu.
\end{aligned}
$$

Then, it follows that

$$
\begin{aligned}
\int_{\mathbb{R}^n} I_1(u; u, u) dx \leq\ & C\|a^u\|_{L^\infty} \|\{\mathbf{I}-\mathbf{P}\}u\|^2_\nu \\
&+ C\|(a^u, b^u)\|_{L^2_x} \|b^u\|_{L^\infty_x} \|\{\mathbf{I}-\mathbf{P}\}u\|_\nu \\
&+ \frac{1}{2}\|b^u\|_{L^\infty_x} \|\{\mathbf{I}-\mathbf{P}\}u\|^2_\nu. \qquad (A.5)
\end{aligned}
$$

For $I_2(u; u, u)$ one has

$$
\begin{aligned}
\int_{\mathbb{R}^n} I_2(u; u, u) dx &= \int_{\mathbb{R}^n} -U * a^u |b^u|^2 + U * b^u \cdot a^u b^u dx \\
&= \iint_{\mathbb{R}^n \times \mathbb{R}^n} U(|x-y|) a^u(x) b^u(y) (b^u(x) - b^u(y)) dx dy \\
&= -\iint_{\mathbb{R}^n \times \mathbb{R}^n} U(|x-y|) a^u(x) (b^u(x) - b^u(y))^2 dx dy \\
&\quad + \iint_{\mathbb{R}^n \times \mathbb{R}^n} U(|x-y|) a^u(x) b^u(x) (b^u(x) - b^u(y)) dx dy.
\end{aligned}
$$

Since

$$
-\iint_{\mathbb{R}^n \times \mathbb{R}^n} U(|x-y|) a^u(x) (b^u(x) - b^u(y))^2 dx dy \leq \|a^u\|_{L^\infty_x} \|T_\Delta b^u\|^2_U,
$$

and

$$
\begin{aligned}
&\iint_{\mathbb{R}^n \times \mathbb{R}^n} U(|x-y|) a^u(x) b^u(x) (b^u(x) - b^u(y)) dx dy \\
&\leq \left[\iint_{\mathbb{R}^n \times \mathbb{R}^n} U(|x-y|)^2 |a^u(x) b^u(x)|^2 dx dy\right]^{\frac{1}{2}} \|T_\Delta b^u\|_U \\
&\leq \|a^u b^u\|_{L^2_x} \|T_\Delta b^u\|_U,
\end{aligned}
$$

it follows that

$$
\int_{\mathbb{R}^n} I_2(u; u, u) dx \leq \|a^u\|_{L^\infty_x} \|T_\Delta b^u\|^2_U + \|a^u b^u\|_{L^2_x} \|T_\Delta b^u\|_U. \qquad (A.6)
$$

Plugging estimates (A.5) and (A.6) into (A.3) and using the coercivity inequality (2.9) of $-\mathbf{L}$, we obtain (4.2). This completes the proof of Lemma 4.1.

Proof of Lemma 4.2: Let $1 \leq |\alpha| \leq N$. Notice that although $\mathbf{A}$ is nonlocal in x,

$$
\partial^\alpha_x \mathbf{A} u = \partial^\alpha_x (U * \rho_{\xi\sqrt{\mathbf{M}}u} \cdot \xi\sqrt{\mathbf{M}}) = U * \rho_{\xi\sqrt{\mathbf{M}}\partial^\alpha_x u} \cdot \xi\sqrt{\mathbf{M}} = \mathbf{A}\partial^\alpha_x u \qquad (A.7)
$$

and hence

$$\partial_x^\alpha \mathbf{L} u = \partial_x^\alpha \mathbf{L}_{FP} u + \partial_x^\alpha \mathbf{A} u = \mathbf{L}_{FP} \partial_x^\alpha u + \mathbf{A} \partial_x^\alpha u = \mathbf{L} \partial_x^\alpha u$$

holds. Then, from (2.12), the energy estimate on $\partial_x^\alpha u$ over $\mathbb{R}_x^n \times \mathbb{R}_\xi^n$ gives

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|\partial_x^\alpha u(t)\|^2 - \int_{\mathbb{R}^n} \langle \mathbf{L}\partial_x^\alpha u, \partial_x^\alpha u\rangle &= \int_{\mathbb{R}^n} \langle \partial_x^\alpha \Gamma(u,u), \partial_x^\alpha u\rangle dx \\ &+ \sum_{\alpha'<\alpha} C_{\alpha'}^\alpha \int_{\mathbb{R}^n} \langle -U * \partial_x^{\alpha-\alpha'} b^u \cdot \nabla_\xi \partial_x^{\alpha'} u, \partial_x^\alpha u\rangle dx. \end{aligned} \tag{A.8}$$

Next we estimate each term on the r.h.s. of (A.8). Similar to (A.7), from Lemma A.1 we obtain

$$\int_{\mathbb{R}^n} \langle \partial_x^\alpha \Gamma(u,u), \partial_x^\alpha u\rangle dx = \sum_{\alpha'\leq\alpha} C_{\alpha'}^\alpha \int_{\mathbb{R}^n} \langle \Gamma(\partial_x^{\alpha'} u, \partial_x^{\alpha-\alpha'} u), \partial_x^\alpha u\rangle dx \leq C I_{\alpha,\alpha'},$$

where

$$\begin{aligned} I_{\alpha,\alpha'} = \int_{\mathbb{R}^n} |U * \partial_x^{\alpha'}(a^u, b^u)| \cdot (|\{\mathbf{I}-\mathbf{P}\}\partial_x^{\alpha-\alpha'} u|_\nu + |\partial_x^{\alpha-\alpha'}(a^u,b^u)|) \\ \cdot (|\{\mathbf{I}-\mathbf{P}\}\partial_x^{\alpha} u|_\nu + |\partial_x^{\alpha}(a^u,b^u)|) dx. \end{aligned}$$

When $|\alpha'| \geq [n/2]+1$ or $\alpha' = \alpha$, one has

$$\begin{aligned} I_{\alpha,\alpha'} \leq \|U * \partial_x^{\alpha'}(a^u,b^u)\|_{L_x^2} (\|\{\mathbf{I}-\mathbf{P}\}\partial_x^\alpha u\|_\nu + \|\partial_x^\alpha (a^u,b^u)\|_{L_x^2}) \\ \times (\sup_x |\{\mathbf{I}-\mathbf{P}\}\partial_x^{\alpha-\alpha'} u|_\nu + \sup_x |\partial_x^{\alpha-\alpha'}(a^u,b^u)|). \end{aligned}$$

By the Sobolev inequality $\|g\|_{L_x^\infty} \leq C\|\nabla_x g\|_{H_x^{[n/2]}}$ for $n \geq 3$ and $g = g(x)$, it further holds that

$$\begin{aligned} &\sup_x |\{\mathbf{I}-\mathbf{P}\}\partial_x^{\alpha-\alpha'} u|_\nu^2 \\ &= \sup_x \int_{\mathbb{R}^n} |\nabla_\xi \{\mathbf{I}-\mathbf{P}\}\partial_x^{\alpha-\alpha'} u|^2 + \nu(\xi)|\{\mathbf{I}-\mathbf{P}\}\partial_x^{\alpha-\alpha'} u|^2 d\xi \\ &\leq \int_{\mathbb{R}^n} \|\nabla_\xi \{\mathbf{I}-\mathbf{P}\}\partial_x^{\alpha-\alpha'} u\|_{L_x^\infty}^2 + \nu(\xi)\|\{\mathbf{I}-\mathbf{P}\}\partial_x^{\alpha-\alpha'} u\|_{L_x^\infty}^2 d\xi \\ &\leq C \int_{\mathbb{R}^n} \|\nabla_\xi \nabla_x \{\mathbf{I}-\mathbf{P}\}\partial_x^{\alpha-\alpha'} u\|_{H_x^{[n/2]}}^2 + \nu(\xi)\|\{\mathbf{I}-\mathbf{P}\}\nabla_x \partial_x^{\alpha-\alpha'} u\|_{H_x^{[n/2]}}^2 d\xi \\ &\leq C \sum_{1\leq|\alpha'|\leq N} \|\{\mathbf{I}-\mathbf{P}\}\partial_x^{\alpha'} u\|_\nu^2, \end{aligned}$$

and similarly

$$\sup_x |\partial_x^{\alpha-\alpha'}(a^u,b^u)| \leq C\|\nabla_x \partial_x^{\alpha-\alpha'}(a^u,b^u)\|_{H_x^{[n/2]}} \leq C\|\nabla_x(a^u,b^u)\|_{H_x^{N-1}}.$$

Hence it follows that

$$I_{\alpha,\alpha'} \leq C\|\nabla_x(a^u,b^u)\|_{H_x^{N-1}} \left(\sum_{1\leq|\alpha'|\leq N} \|\partial_x^{\alpha'}\{\mathbf{I}-\mathbf{P}\}u\|_\nu^2 + \|\nabla_x(a^u,b^u)\|_{H_x^{N-1}}^2 \right). \tag{A.9}$$

When $|\alpha'| \leq [n/2]$ and $\alpha' < \alpha$, one obtains similarly that

$$\begin{aligned} I_{\alpha,\alpha'} \leq \sup_x |U * \partial_x^{\alpha'}(a^u,b^u)| \cdot (\|\{\mathbf{I}-\mathbf{P}\}\partial_x^{\alpha-\alpha'} u\|_\nu + \|\partial_x^{\alpha-\alpha'}(a^u,b^u)\|_{L_x^2}) \\ \cdot (\|\{\mathbf{I}-\mathbf{P}\}\partial_x^{\alpha} u\|_\nu + \|\partial_x^{\alpha}(a^u,b^u)\|_{L_x^2}). \end{aligned}$$

A further bound also holds for the r.h.s. of (A.9) since

$$\sup_x |U * \partial_x^{\alpha'}(a^u, b^u)| \leq \sup_x |\partial_x^{\alpha'}(a^u, b^u)| \leq C\|\nabla_x(a^u, b^u)\|_{H_x^{N-1}}.$$

For the second term on the r.h.s. of (A.8), by using the same proof as for $I_{\alpha,\alpha'}$, one has

$$\begin{aligned}
&\sum_{\alpha'<\alpha} C_{\alpha'}^{\alpha} \int_{\mathbb{R}^n} \langle -U * \partial_x^{\alpha-\alpha'} b^u \cdot \nabla_\xi \partial_x^{\alpha'} u, \partial_x^\alpha u\rangle dx \\
&\leq C \sum_{\alpha'<\alpha} \int_{\mathbb{R}^n} |U * \partial_x^{\alpha-\alpha'} b^u| \cdot \|\nabla_\xi \partial_x^{\alpha'} u\|_{L_\xi^2} \|\partial_x^\alpha u\|_{L_\xi^2} dx \\
&\leq C\|\nabla_x b^u\|_{H_x^{N-1}} \left(\sum_{1\leq|\alpha'|\leq N-1} \|\nabla_\xi \partial_x^{\alpha'} u\|^2 + \|\partial_x^\alpha u\|^2 \right) \\
&\leq C\|\nabla_x b^u\|_{H_x^{N-1}} \sum_{1\leq|\alpha'|\leq N-1} \|\partial_x^{\alpha'} \nabla_\xi \{\mathbf{I}-\mathbf{P}\}u\|^2 \\
&\quad + C\|\nabla_x b^u\|_{H_x^{N-1}} \left(\sum_{1\leq|\alpha'|\leq N} \|\partial_x^{\alpha'} \{\mathbf{I}-\mathbf{P}\}u\|_\nu^2 + \|\nabla_x(a^u, b^u)\|_{H_x^{N-1}}^2 \right).
\end{aligned}$$

Putting all the above estimates into (A.8), taking summation over $1 \leq |\alpha| \leq N$ and using the coercivity inequality (2.9) of $-\mathbf{L}$ yields (4.3). This completes the proof of Lemma 4.2.

Proof of Lemma 4.3: Fix k with $1 \leq k \leq N$, and choose α, β with $|\beta| = k$ and $|\alpha| + |\beta| \leq N$. From (4.5), the energy estimate on $\partial_x^\alpha \partial_\xi^\beta u$ over $\mathbb{R}_x^n \times \mathbb{R}_\xi^n$ gives

$$\frac{1}{2}\frac{d}{dt}\|\partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u\|^2 - \int_{\mathbb{R}^n} \langle \mathbf{L}_{FP} \partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u, \partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u\rangle dx = \sum_{i=1}^{7} I_i, \tag{A.10}$$

where I_i $(1 \leq i \leq 7)$ take the form of

$$\begin{aligned}
I_1 &= \int_{\mathbb{R}^n} \langle -\partial_x^\alpha [\partial_\xi^\beta, \xi \cdot \nabla_x]\{\mathbf{I}-\mathbf{P}\}u, \partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u\rangle dx, \\
I_2 &= \int_{\mathbb{R}^n} \langle \partial_x^\alpha [\partial_\xi^\beta, -|\xi|^2]\{\mathbf{I}-\mathbf{P}\}u, \partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u\rangle dx,
\end{aligned}$$

and

$$\begin{aligned}
I_3 &= \sum_{\alpha'<\alpha} C_{\alpha'}^{\alpha} \int_{\mathbb{R}^n} \langle -U * \partial_x^{\alpha-\alpha'} b^u \cdot \nabla_\xi \partial_x^{\alpha'} \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u, \partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u\rangle dx, \\
I_4 &= \int_{\mathbb{R}^n} \langle \partial_x^\alpha \partial_\xi^\beta \Gamma(u, \{\mathbf{I}-\mathbf{P}\}u), \partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u\rangle dx, \\
I_5 &= \int_{\mathbb{R}^n} \langle \frac{1}{2}\partial_x^\alpha \partial_\xi^\beta (U * b^u \cdot [\xi, \mathbf{P}]u), \partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u\rangle dx,
\end{aligned}$$

and

$$\begin{aligned}
I_6 &= \int_{\mathbb{R}^n} \langle \partial_x^\alpha \partial_\xi^\beta \mathbf{P}(\xi \cdot \nabla_x \{\mathbf{I}-\mathbf{P}\}u + U * b^u \cdot \nabla_\xi \{\mathbf{I}-\mathbf{P}\}u), \partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u\rangle dx, \\
I_7 &= \int_{\mathbb{R}^n} \langle -\partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}(\xi \cdot \nabla_x \mathbf{P}u + U * b^u \cdot \nabla_\xi \mathbf{P}u), \partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u\rangle dx.
\end{aligned}$$

Here, the commutator in I_2 follows from

$$[\partial_\xi^\beta, \mathbf{L}_{FP}] = [\partial_\xi^\beta, \Delta_\xi] + [\partial_\xi^\beta, \frac{1}{4}(2n - |\xi|^2)] = [\partial_\xi^\beta, -|\xi|^2]. \tag{A.11}$$

Next, we estimate each term I_i $(1 \leq i \leq 7)$. For I_1 and I_2, one has

$$\begin{aligned} I_1 &\leq \delta\|\partial_x^\alpha\partial_\xi^\beta\{\mathbf{I}-\mathbf{P}\}u\|^2 + C_\delta\|[\partial_\xi^\beta, \xi\cdot\nabla_x]\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u\|^2 \\ &\leq \delta\|\partial_x^\alpha\partial_\xi^\beta\{\mathbf{I}-\mathbf{P}\}u\|^2 + C_\delta\sum_{|\alpha|\leq N-k}\|\partial_x^\alpha\nabla_x\{\mathbf{I}-\mathbf{P}\}u\|^2 \\ &\quad +\chi_{\{2\leq k\leq N\}}C_\delta\sum_{\substack{1\leq|\beta|\leq k-1\\ |\alpha|+|\beta|\leq N}}\|\partial_x^\alpha\partial_\xi^\beta\{\mathbf{I}-\mathbf{P}\}u\|^2, \end{aligned}$$

and

$$\begin{aligned} I_2 &\leq \delta\|\partial_x^\alpha\partial_\xi^\beta\{\mathbf{I}-\mathbf{P}\}u\|^2 + C_\delta\|[\partial_\xi^\beta, -|\xi|^2]\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u\|^2 \\ &\leq \delta\|\partial_x^\alpha\partial_\xi^\beta\{\mathbf{I}-\mathbf{P}\}u\|^2 + C_\delta\sum_{|\alpha|\leq N-k}\|\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u\|_\nu^2 \\ &\quad +\chi_{\{2\leq k\leq N\}}C_\delta\sum_{\substack{1\leq|\beta|\leq k-1\\ |\alpha|+|\beta|\leq N}}\|\partial_x^\alpha\partial_\xi^\beta\{\mathbf{I}-\mathbf{P}\}u\|_\nu^2, \end{aligned}$$

where $\delta > 0$ is arbitrary, to be chosen later. For I_3 and I_5, similar to the proof of (A.9), it holds

$$\begin{aligned} I_3 &\leq C\sum_{\alpha'<\alpha}\int_{\mathbb{R}^n}|U*\partial_x^{\alpha-\alpha'}b^u|\cdot\|\nabla_\xi\partial_x^{\alpha'}\partial_\xi^\beta\{\mathbf{I}-\mathbf{P}\}u\|_{L^2_\xi}\|\partial_x^\alpha\partial_\xi^\beta\{\mathbf{I}-\mathbf{P}\}u\|_{L^2_\xi}dx \\ &\leq C\|\nabla_x b^u\|_{H_x^{N-1}}\sum_{|\alpha'|+|\beta'|\leq N}\|\partial_x^{\alpha'}\partial_\xi^{\beta'}\{\mathbf{I}-\mathbf{P}\}u\|^2, \end{aligned}$$

and

$$\begin{aligned} I_5 &\leq C\sum_{\alpha'\leq\alpha}\int_{\mathbb{R}^n}|U*\partial_x^{\alpha-\alpha'}b^u|\cdot\|\partial_\xi^\beta([\xi,\mathbf{P}]\partial_x^{\alpha'}u)\|_{L^2_\xi}\|\partial_x^\alpha\partial_\xi^\beta\{\mathbf{I}-\mathbf{P}\}u\|_{L^2_\xi}dx \\ &\leq C\sum_{\alpha'\leq\alpha}\int_{\mathbb{R}^n}|U*\partial_x^{\alpha-\alpha'}b^u|\cdot|\partial_x^{\alpha'}(a^u,b^u)|\cdot\|\partial_x^\alpha\partial_\xi^\beta\{\mathbf{I}-\mathbf{P}\}u\|_{L^2_\xi}dx \\ &\leq C\|(a^u,b^u)\|_{H_x^N}(\|\nabla_x(a^u,b^u)\|^2_{H_x^{N-1}} + \|\partial_x^\alpha\partial_\xi^\beta\{\mathbf{I}-\mathbf{P}\}u\|^2). \end{aligned}$$

For I_6 and I_7, it follows in the same way that

$$\begin{aligned} I_6 &\leq \delta\|\partial_x^\alpha\partial_\xi^\beta\{\mathbf{I}-\mathbf{P}\}u\|^2 + C_\delta\sum_{|\alpha'|\leq N-k}\|\nabla_x\partial_x^{\alpha'}\{\mathbf{I}-\mathbf{P}\}u\|^2 \\ &\quad +C\|b^u\|_{H_x^N}\sum_{|\alpha'|+|\beta'|\leq N}\|\partial_x^{\alpha'}\partial_\xi^{\beta'}\{\mathbf{I}-\mathbf{P}\}u\|^2, \end{aligned}$$

and

$$\begin{aligned} I_7 &\leq \delta\|\partial_x^\alpha\partial_\xi^\beta\{\mathbf{I}-\mathbf{P}\}u\|^2 + C_\delta\|\nabla_x(a^u,b^u)\|^2_{H_x^{N-k}} \\ &\quad +C\|(a^u,b^u)\|_{H_x^N}(\|\nabla_x(a^u,b^u)\|^2_{H_x^{N-1}} + \|\partial_x^\alpha\partial_\xi^\beta\{\mathbf{I}-\mathbf{P}\}u\|^2). \end{aligned}$$

Now, it remains to estimate I_4. To this extent, one can use the identity

$$\begin{aligned} \partial_\xi^\beta\Gamma(u,v) &= \Gamma(u,\partial_\xi^\beta v) + U*a^u[\partial_\xi^\beta, \mathbf{L}_{FP}]v + \frac{1}{2}U*b^u\cdot[\partial_\xi^\beta,\xi]v \\ &= \Gamma(u,\partial_\xi^\beta v) + U*a^u[\partial_\xi^\beta, -|\xi|^2]v + \frac{1}{2}U*b^u\cdot[\partial_\xi^\beta,\xi]v, \end{aligned}$$

where (A.11) was used again. Then, I_4 can be rewritten as

$$I_4 = I_{4,1} + I_{4,2},$$

with

$$I_{4,1} = \int_{\mathbb{R}^n} \langle \partial_x^\alpha \Gamma(u, \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u), \partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u\rangle dx,$$

$$I_{4,2} = \int_{\mathbb{R}^n} \langle \partial_x^\alpha (U * a^u[\partial_\xi^\beta, -|\xi|^2]\{\mathbf{I}-\mathbf{P}\}u), \partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u\rangle dx + \int_{\mathbb{R}^n} \langle \partial_x^\alpha (\frac{1}{2} U * b^u \cdot [\partial_\xi^\beta, \xi]\{\mathbf{I}-\mathbf{P}\}u), \partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u\rangle dx.$$

Similarly to the bound on I_3, $I_{4,2}$ can be controlled by

$$I_{4,2} \leq C\|(a^u, b^u)\|_{H_x^N} \sum_{|\alpha'|+|\beta'|\leq N} \|\partial_x^{\alpha'} \partial_\xi^{\beta'} \{\mathbf{I}-\mathbf{P}\}u\|_\nu^2.$$

For $I_{4,1}$, it follows from Lemma A.1 that

$$\begin{aligned} I_{4,1} &= \sum_{\alpha'\leq\alpha} C_{\alpha'}^\alpha \int_{\mathbb{R}^n} \langle \Gamma(\partial_x^{\alpha-\alpha'} u, \partial_x^{\alpha'} \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u), \partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u\rangle dx \\ &\leq C \sum_{\alpha'\leq\alpha} \int_{\mathbb{R}^n} |U * (a(\partial_x^{\alpha-\alpha'} u), b(\partial_x^{\alpha-\alpha'} u))| \\ &\quad \times [|\{\mathbf{I}-\mathbf{P}\}\partial_x^{\alpha'} \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u|_\nu + |(a(\partial_x^{\alpha'} \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u), b(\partial_x^{\alpha'} \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u))|] \\ &\quad \times [|\{\mathbf{I}-\mathbf{P}\}\partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u|_\nu + |(a(\partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u), b(\partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u))|]dx, \end{aligned}$$

where $a(w) = a^w$, $b(w) = b^w$ for any w. For any α', β', one can use the properties

$$a(\partial_x^{\alpha'} u) = \partial_x^{\alpha'} a^u, \quad |a(\partial_x^{\alpha'} \partial_\xi^{\beta'} \{\mathbf{I}-\mathbf{P}\}u)| \leq C\|\partial_x^{\alpha'} \{\mathbf{I}-\mathbf{P}\}u\|_{L_\xi^2},$$

and similarly for b. Then, it further holds that

$$\begin{aligned} I_{4,2} &\leq C \sum_{\alpha'\leq\alpha} \int_{\mathbb{R}^n} |U * \partial_x^{\alpha-\alpha'} (a^u, b^u)| \\ &\qquad \times (|\{\mathbf{I}-\mathbf{P}\}\partial_x^{\alpha'} \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u|_\nu + \|\partial_x^{\alpha'} \{\mathbf{I}-\mathbf{P}\}u\|_{L_\xi^2}) \\ &\qquad \times (|\{\mathbf{I}-\mathbf{P}\}\partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u|_\nu + \|\partial_x^\alpha \{\mathbf{I}-\mathbf{P}\}u\|_{L_\xi^2})dx \\ &\leq C\|(a^u, b^u)\|_{H_x^N} \sum_{|\alpha'|+|\beta'|\leq N} \|\partial_x^{\alpha'} \partial_\xi^{\beta'} \{\mathbf{I}-\mathbf{P}\}u\|_\nu^2. \end{aligned}$$

Finally, using the coercivity inequality (2.4) of $-\mathbf{L}_{FP}$, the second term on the l.h.s. of (A.10)satisfies the low bound

$$\begin{aligned} & -\int_{\mathbb{R}^n} \langle \mathbf{L}_{FP} \partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u, \partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u\rangle dx \\ & \geq \lambda_0 \|\{\mathbf{I}-\mathbf{P}_0\}\partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u\|_\nu^2 \\ & \geq \frac{\lambda_0}{2} \|\partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u\|_\nu^2 - \lambda_0 \|\mathbf{P}_0 \partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u\|_\nu^2 \\ & \geq \frac{\lambda_0}{2} \|\partial_x^\alpha \partial_\xi^\beta \{\mathbf{I}-\mathbf{P}\}u\|_\nu^2 - C\|\partial_x^\alpha \{\mathbf{I}-\mathbf{P}\}u\|^2. \end{aligned}$$

Finally, plugging all the above estimates into (A.10), taking summation over $\{|\beta| = k, |\alpha| + |\beta| \leq N\}$ and then choosing a properly small $\delta > 0$ yields (4.6). This completes the proof of Lemma 4.3.

A.2 A priori estimates: Macroscopic dissipation

In this second part of the appendix, we shall prove Lemma 4.4 for the macroscopic dissipation which plays a key role in the proof of global existence. We first prove estimates on the space derivatives of $A_{ij}(l)$ and $A_{ij}(r)$ in the following

Lemma A.2. *Let l, r be given by* (2.17) *and* (2.18) *respectively. it holds*

$$\sum_{|\alpha|\le N-1}\|\partial_x^\alpha A_{ij}(l)\|_{L_x^2}\le C\sum_{|\alpha|\le N}\|\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u\|^2, \tag{A.12}$$

$$\sum_{|\alpha|\le N-1}\|\partial_x^\alpha A_{ij}(r)\|_{L_x^2}\le C\|\nabla_x(a^u,b^u)\|_{H_x^{N-1}}\sum_{|\alpha|\le N}\|\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u\|, \tag{A.13}$$

for some constant C depending only on n, where the moment function $A_{ij}(\cdot)$, $1\le i,j\le n$, are defined by (2.20)*.*

Proof. For any α, integration by parts yields straightforwardly

$$\begin{aligned}
&|\partial_x^\alpha A_{ij}(l)| = |A_{ij}(\partial_x^\alpha l)|\\
&= |\langle(\xi_i\xi_j-1)\sqrt{\mathbf{M}}, -\xi\cdot\nabla_x\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u+\mathbf{L}_{FP}\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u\rangle|\\
&\le |\langle-\xi(\xi_i\xi_j-1)\sqrt{\mathbf{M}}, \nabla_x\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u\rangle|+|\langle\mathbf{L}_{FP}([\xi_i\xi_j-1]\sqrt{\mathbf{M}}), \partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u\rangle|\\
&\le \|\xi(\xi_i\xi_j-1)\sqrt{\mathbf{M}}\|_{L_\xi^2}\|\nabla_x\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u\|_{L_\xi^2}\\
&\qquad\qquad +\|\mathbf{L}_{FP}([\xi_i\xi_j-1]\sqrt{\mathbf{M}})\|_{L_\xi^2}\|\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u\|_{L_\xi^2}\\
&\le C(\|\nabla_x\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u\|_{L_\xi^2}+\|\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u\|_{L_\xi^2}).
\end{aligned}$$

The exponential decay in ξ for $\mathbf{M}$ was used above. Thus, (A.12) follows by further taking L_x^2-norm and then summation over $|\alpha|\le N-1$. For $A_{ij}(n)$, one has

$$\begin{aligned}
&\partial_x^\alpha A_{ij}(r) = A_{ij}(\partial_x^\alpha r)\\
&= \langle(\xi_i\xi_j-1)\sqrt{\mathbf{M}}, \partial_x^\alpha[U*a^u\mathbf{L}_{FP}\{\mathbf{I}-\mathbf{P}\}u]\\
&\qquad\qquad +\frac{1}{2}U*b^u\cdot\xi\{\mathbf{I}-\mathbf{P}\}u-U*b^u\cdot\nabla_\xi\{\mathbf{I}-\mathbf{P}\}u]\rangle\\
&= \sum_{\alpha'\le\alpha}C_{\alpha'}^\alpha\langle(\xi_i\xi_j-1)\sqrt{\mathbf{M}}, U*\partial_x^{\alpha-\alpha'}a^u\mathbf{L}_{FP}\partial_x^{\alpha'}\{\mathbf{I}-\mathbf{P}\}u\\
&\qquad +\frac{1}{2}U*\partial_x^{\alpha-\alpha'}b^u\cdot\xi\partial_x^{\alpha'}\{\mathbf{I}-\mathbf{P}\}u-U*\partial_x^{\alpha-\alpha'}b^u\cdot\nabla_\xi\partial_x^{\alpha'}\{\mathbf{I}-\mathbf{P}\}u\rangle,
\end{aligned}$$

which from integration by parts gives

$$\begin{aligned}
\partial_x^\alpha A_{ij}(r) = \sum_{\alpha'\le\alpha}C_{\alpha'}^\alpha\Big\{&\langle\mathbf{L}_{FP}([\xi_i\xi_j-1]\sqrt{\mathbf{M}}), U*\partial_x^{\alpha-\alpha'}a^u\partial_x^{\alpha'}\{\mathbf{I}-\mathbf{P}\}u\rangle\\
&+\langle\frac{1}{2}\xi(\xi_i\xi_j-1)\sqrt{\mathbf{M}}, U*\partial_x^{\alpha-\alpha'}b^u\partial_x^{\alpha'}\{\mathbf{I}-\mathbf{P}\}u\rangle\\
&+\langle\nabla_\xi((\xi_i\xi_j-1)\sqrt{\mathbf{M}})\rangle, U*\partial_x^{\alpha-\alpha'}b^u\partial_x^{\alpha'}\{\mathbf{I}-\mathbf{P}\}u\rangle\Big\}.
\end{aligned}$$

Thus, it follows from the Hölder inequality that

$$|\partial_x^\alpha A_{ij}(r)|\le C\sum_{\alpha'\le\alpha}|U*\partial_x^{\alpha'-\alpha}(a^u,b^u)|\cdot\|\partial_x^{\alpha'}\{\mathbf{I}-\mathbf{P}\}u\|_{L_\xi^2}$$

for any α. Hence, similarly as before, (A.13) follows by taking further L_x^2-norm, applying the Young and Sobolev inequalities and then taking summation over $|\alpha|\le N-1$. This completes the proof of Lemma A.2. □

Proof of Lemma 4.4: Recall the equations (2.21)-(2.24) satisfied by (a^u, b^u) and the parabolic-type equation (2.25) derived from (2.23)-(2.24). We begin with the estimate on b^u from (2.22) and (2.25). Let

$|\alpha| \leq N-1$. It follows from (2.25) that

$$\begin{aligned}
&\frac{d}{dt}\int_{\mathbb{R}^n} \partial_x^\alpha[\sum_{i\neq j}\partial_j A_{ii}(\{\mathbf{I}-\mathbf{P}\}u) - \sum_i \partial_i A_{ij}(\{\mathbf{I}-\mathbf{P}\}u)]\partial_x^\alpha b_j^u dx + \int_{\mathbb{R}^n} |\nabla_x \partial_x^\alpha b_j^u|^2 dx\\
&= \int_{\mathbb{R}^n} \partial_x^\alpha[\sum_{i\neq j}\partial_j A_{ii}(\{\mathbf{I}-\mathbf{P}\}u) - \sum_i \partial_i A_{ij}(\{\mathbf{I}-\mathbf{P}\}u)]\partial_x^\alpha \partial_t b_j^u dx\\
&\quad + \int_{\mathbb{R}^n} \partial_x^\alpha[\sum_i \partial_j(U * b_i^u b_i^u) - \sum_i \partial_i(U * b_i^u b_j^u + U * b_j^u b_i^u)]\partial_x^\alpha b_j^u dx\\
&\quad + \int_{\mathbb{R}^n} \partial_x^\alpha[\sum_{i\neq j}\partial_j A_{ii}(l+n) - \sum_i \partial_i A_{ij}(l+n)]\partial_x^\alpha b_j^u dx\\
&=: I_1 + I_2 + I_3, \qquad \text{(A.14)}
\end{aligned}$$

where I_i, $i=1,2,3$, denote the corresponding terms on the r.h.s., respectively. For I_3, it holds

$$I_3 \leq \lambda\|\nabla_x \partial_x^\alpha b_j^u\|_{L_x^2}^2 + \frac{C}{\lambda}\sum_{ij}(\|\partial_x^\alpha A_{ij}(l)\|_{L_x^2}^2 + \|\partial_x^\alpha A_{ij}(n)\|_{L_x^2}^2),$$

where $0<\lambda\leq 1$ is a constant to be chosen later. For I_2, similarly one has

$$\begin{aligned}
I_2 &\leq \lambda\|\nabla_x \partial_x^\alpha b_j^u\|_{L_x^2}^2 + \frac{C}{\lambda}\sum_{ij}\|\partial_x^\alpha(U * b_i^u b_j^u)\|_{L_x^2}^2\\
&\leq \lambda\|\nabla_x \partial_x^\alpha b_j^u\|_{L_x^2}^2 + \frac{C}{\lambda}\sum_{ij}\sum_{\alpha'\leq\alpha}\|U * \partial_x^{\alpha-\alpha'} b_i^u \partial_x^{\alpha'} b_j^u\|_{L_x^2}^2\\
&\leq \lambda\|\nabla_x \partial_x^\alpha b_j^u\|_{L_x^2}^2 + \frac{C}{\lambda}\|b^u\|_{H_x^N}^2\|\nabla_x b^u\|_{H_x^{N-1}}^2,
\end{aligned}$$

where the Young and Sobolev inequalities were used. For I_1, one can use (2.22) to rewrite it as

$$\begin{aligned}
I_1 = \int_{\mathbb{R}^n} &\partial_x^\alpha[\sum_{i\neq j}\partial_j A_{ii}(\{\mathbf{I}-\mathbf{P}\}u) - \sum_i \partial_i A_{ij}(\{\mathbf{I}-\mathbf{P}\}u)]\\
&\cdot\partial_x^\alpha[-\partial_i a^u + (U * b_i^u - b_i^u) - (U * a^u b_i^u - U * b_i^u a^u)\\
&\qquad + \sum_j \partial_j A_{ij}(\{\mathbf{I}-\mathbf{P}\}u)]dx.
\end{aligned}$$

Then, it holds

$$\begin{aligned}
I_1 &\leq \delta\|\nabla_x \partial_x^\alpha a_j^u\|_{L_x^2}^2 + \frac{C}{\delta}\sum_{ij}\|A_{ij}(\nabla_x \partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u)\|_{L_x^2}^2\\
&+\lambda\|\nabla_x \partial_x^\alpha b_j^u\|_{L_x^2}^2 + \frac{C}{\lambda}\sum_{ij}\|A_{ij}(\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u)\|_{L_x^2}^2\\
&+\|\partial_x^\alpha(U * a^u b_i^u - U * b_i^u a^u)\|_{L_x^2}^2 + C\sum_{ij}\|A_{ij}(\nabla_x \partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u)\|_{L_x^2}^2,
\end{aligned}$$

where $0<\delta\leq 1$ is a constant to be also determined later, and again from the Young and Sobolev inequalities, one has

$$\|\partial_x^\alpha(U * a^u b_i^u - U * b_i^u a^u)\|_{L_x^2} \leq C\|(a^u,b^u)\|_{H_x^N}\|\nabla_x(a^u,b^u)\|_{H_x^{N-1}}.$$

Thus, it follows that

$$\begin{aligned}
I_1 &\leq \lambda\|\nabla_x \partial_x^\alpha b_j^u\|_{L_x^2}^2 + \delta\|\nabla_x \partial_x^\alpha a_j^u\|_{L_x^2}^2\\
&+C\|(a^u,b^u)\|_{H_x^N}^2\|\nabla_x(a^u,b^u)\|_{H_x^{N-1}}^2 + C_{\lambda,\delta}\sum_{|\alpha|\leq N}\|\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u\|^2.
\end{aligned}$$

By plugging the above estimates on I_1, I_2 and I_3 into (A.14), taking summation over $|\alpha| \leq N-1$, $1 \leq j \leq n$ and then choosing $0 < \lambda \leq 1$ properly small, one has

$$\begin{aligned}
&\frac{d}{dt}\mathcal{E}^{n,b}_{free}(u(t)) + \frac{1}{2}\|\nabla_x b^u\|^2_{H^{N-1}_x}\\
&\leq \delta\|\nabla_x a^u\|^2_{H^{N-1}_x} + C_\delta \sum_{|\alpha|\leq N}\|\partial^\alpha_x\{\mathbf{I}-\mathbf{P}\}u\|^2\\
&\quad +C\|(a^u,b^u)\|^2_{H^N_x}\|\nabla_x(a^u,b^u)\|^2_{H^{N-1}_x}\\
&\quad +C\sum_{ij}(\|\partial^\alpha_x A_{ij}(l)\|^2_{L^2_x} + \|\partial^\alpha_x A_{ij}(n)\|^2_{L^2_x}), \qquad \text{(A.15)}
\end{aligned}$$

where the free energy $\mathcal{E}^{n,b}_{free}(u(t))$ corresponding to the dissipation of b^u is given by

$$\begin{aligned}
\mathcal{E}^{n,b}_{free}(u(t)) = &\sum_{|\alpha|\leq N-1}\sum_j\sum_{i\neq j}\int_{\mathbb{R}^n} A_{ii}(\partial^\alpha_x\partial_j\{\mathbf{I}-\mathbf{P}\}u)\partial^\alpha_x b^u_j dx\\
&-\sum_{|\alpha|\leq N-1}\sum_{ij}\int_{\mathbb{R}^n} A_{ij}(\partial^\alpha_x\partial_i\{\mathbf{I}-\mathbf{P}\}u)\partial^\alpha_x b^u_j dx.
\end{aligned}$$

Next, we shall obtain the dissipation of a^u from (2.21) and (2.22). Let $|\alpha| \leq N-1$. It follows from (2.22) that

$$\begin{aligned}
&\frac{d}{dt}\int_{\mathbb{R}^n}\partial^\alpha_x\nabla_x a^u\cdot\partial^\alpha_x b^u dx + \|\nabla_x\partial^\alpha_x a^u\|^2_{L^2_x}\\
&=\int_{\mathbb{R}^n}\partial^\alpha_x\nabla_x\partial_t a^u\cdot\partial^\alpha_x b^u dx + \int_{\mathbb{R}^n}\partial^\alpha_x\nabla_x a^u\cdot\partial^\alpha_x(U*b^u - b^u)dx\\
&\quad+\int_{\mathbb{R}^n}\partial^\alpha_x a^u\cdot\partial^\alpha_x(U*b^u a^u - U*a^u b^u)dx\\
&\quad-\sum_{ij}\int_{\mathbb{R}^n}\partial^\alpha_x\partial_i a^u\partial^\alpha_x\partial_j A_{ij}(\{\mathbf{I}-\mathbf{P}\}u)dx\\
&= I_4 + I_5 + I_6 + I_7, \qquad \text{(A.16)}
\end{aligned}$$

where I_i, $4 \leq i \leq 7$, denote the corresponding terms on the r.h.s., respectively. From the conservation law of mass (2.21), I_4 is rewritten as

$$I_4 = -\int_{\mathbb{R}^n}\partial^\alpha_x\partial_t a^u\partial^\alpha_x\nabla_x\cdot b^u dx = \|\partial^\alpha_x\nabla_x\cdot b^u\|^2_{L^2_x}.$$

For I_5, one can use the general inequality

$$\begin{aligned}
\left|\int_{\mathbb{R}^n} a(U*b-b)dx\right| &= \left|\iint_{\mathbb{R}^n\times\mathbb{R}^n} a(x)U(|x-y|)(b(y)-b(x))dxdy\right|\\
&\leq\left[\iint_{\mathbb{R}^n\times\mathbb{R}^n}|a(x)|^2U(|x-y|)dxdy\right]^{1/2}\|T_\Delta b\|_U\\
&\leq\|a\|_{L^2_x}\|T_\Delta b\|_U
\end{aligned}$$

for $a = a(x)$ and $b = b(x)$ so that it holds

$$I_5 \leq \frac{1}{6}\|\nabla_x\partial^\alpha_x a^u\|^2_{L^2_x} + C\|T_\Delta\partial^\alpha_x b^u\|^2_U.$$

For I_6 and I_7, similarly as before, one has

$$\begin{aligned}
I_6 &\leq \frac{1}{6}\|\nabla_x\partial^\alpha_x a^u\|^2_{L^2_x} + C\|\partial^\alpha_x(U*b^u a^u - U*a^u b^u)\|^2_{L^2_x}\\
&\leq \frac{1}{6}\|\nabla_x\partial^\alpha_x a^u\|^2_{L^2_x} + C\|(a^u,b^u)\|^2_{H^N_x}\|\nabla_x(a^u,b^u)\|^2_{H^{N-1}_x},
\end{aligned}$$

and

$$
\begin{aligned}
I_7 &\le \frac{1}{6}\|\nabla_x\partial_x^\alpha a^u\|_{L_x^2}^2 + C\sum_{ij}\|A_{ij}(\partial_x^\alpha\partial_j\{\mathbf{I}-\mathbf{P}\}u)\|_{L_x^2}^2\\
&\le \frac{1}{6}\|\nabla_x\partial_x^\alpha a^u\|_{L_x^2}^2 + \|\nabla_x\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u\|^2.
\end{aligned}
$$

Putting the above estimates on I_i $(4\le i\le 7)$ into (A.16) and taking summation over $|\alpha|\le N-1$ gives

$$
\begin{aligned}
&\frac{d}{dt}\mathcal{E}^{n,a}_{free}(u(t)) + \frac{1}{2}\|\nabla_x a^u\|^2_{H_x^{N-1}}\\
&\le \|\nabla_x b^u\|^2_{H_x^{N-1}} + C\sum_{|\alpha|\le N-1}(\|T_\Delta\partial_x^\alpha b^u\|_U^2 + \|\nabla_x\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}u\|^2)\\
&\quad + C\|(a^u,b^u)\|^2_{H_x^N}\|\nabla_x(a^u,b^u)\|^2_{H_x^{N-1}},
\end{aligned}
\tag{A.17}
$$

where the free energy $\mathcal{E}^{n,a}_{free}(u(t))$ corresponding to the dissipation of a^u is given by

$$
\mathcal{E}^{n,a}_{free}(u(t)) = \sum_{|\alpha|\le N-1}\int_{\mathbb{R}^n}\partial_x^\alpha\nabla_x a^u\cdot\partial_x^\alpha b^u dx.
$$

Therefore, (4.7) with $\mathcal{E}^n_{free}(u(t))$ defined by (4.6) follows from taking the proper linear combination of (A.15) and (A.17) and then choosing a properly small $0<\delta\le 1$. Finally, (4.8) holds true since one has

$$
\begin{aligned}
|\mathcal{E}^n_{free}(u(t))| &\le C\sum_{|\alpha|\le N}(\|\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}\|^2 + \|\partial_x^\alpha(a^u,b^u)\|^2_{L_x^2})\\
&\le C\sum_{|\alpha|\le N}(\|\partial_x^\alpha\{\mathbf{I}-\mathbf{P}\}\|^2 + \|\partial_x^\alpha\mathbf{P}\|^2)\\
&\le C\|u(t)\|^2_{L^2_\xi(H_x^N)}.
\end{aligned}
$$

This completes the proof of Lemma 4.4.

Acknowledgments

Renjun Duan and Massimo Fornasier acknowledge the financial support provided by the FWF project Y 432-N15 START-Preis “Sparse Approximation and Optimization in High Dimensions”.

References

[1] 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** (2001), 43–100.

[2] J.A. Carrillo, M. Fornasier, J. Rosado, G. Toscani, Asymptotic flocking dynamics for the kinetic Cucker-Smale model, (preprint) (2009)

[3] Y.-L. Chuang, M. R. D’Orsogna, D. Marthaler, A. L. Bertozzi and L. Chayes, State transitions and the continuum limit for a 2D interacting, self-propelled particle system, *Physica D* **232** (2007), 33–47.

[4] F. Cucker and S. Smale, Emergent behavior in flocks, *IEEE Transactions of Automatic Control* **52** (2007), 852–862.

[5] P. Degond, S. Motsch, Continuum limit of self-driven particles with orientation interaction, *Mathematical Models and Methods in Applied Sciences* **18**, Suppl. (2008), 1193-1215.

[6] P. Degond, S. Motsch, Large scale dynamics of the persistent turning walker model of fish behavior, *J. Stat. Phys.* **131** (2008), 989-1021.

[7] L. Desvillettes and C. Villani, On the trend to global equilibrium in spatially inhomogeneous entropy-dissipating systems: the linear Fokker-Planck equation, *Comm. Pure Appl. Math.* **54** (1) (2001), 1–42.

[8] J. Dolbeault, C. Mouhot and C. Schmeiser, Hypocoercivity for kinetic equations with linear relaxation terms, *C.R. Acad. Sci. Paris* **347** (2009), 511–516.

[9] R.-J. Duan, On the Cauchy problem for the Boltzmann equation in the whole space: Global existence and uniform stability in $L^2_x(H^N_x)$, *Journal of Differential Equations* **244** (2007), 3204–3234.

[10] R.-J. Duan, Stability of the Boltzmann equation with potential forces on Torus, *Physica D: Nonlinear Phenomena*, in press (2009).

[11] R.-J. Duan, S. Ukai and T. Yang, A combination of energy method and spectral analysis for studies on systems for gas motions, *Frontiers of Mathematics in China* **4** (2) (2009), 253–282.

[12] R.-J. Duan, S. Ukai, T. Yang and H.-J. Zhao, Optimal decay estimates on the linearized Boltzmann equation with time-dependent forces and their applications, *Commun. Math. Phys.* **277** (1) (2008), 189–236.

[13] Y. Guo, The Boltzmann equation in the whole space, *Indiana Univ. Math. J.* **53** (2004), 1081–1094.

[14] Y. Guo, The Landau equation in a periodic box, *Commun. Math. Phys.* **231** (2002), 391–434.

[15] S.-Y. Ha and J.-G. Liu, A simple proof of the Cucker-Smale flocking dynamics and mean-field limit, *Commun. Math. Sci.* **7** (2009), 297–325.

[16] S.-Y. Ha and E. Tadmor, From particle to kinetic and hydrodynamic descriptions of flocking, *Kinetic and Related Models* **1** (3) (2008), 415–435.

[17] S. Kawashima, Systems of a hyperbolic-parabolic composite type, with applications to the equations of magnetohydrodynamics, thesis, Kyoto University (1983).

[18] C Mouhot and L. Neumann, Quantitative perturbative study of convergence to equilibrium for collisional kinetic models in the torus, *Nonlinearity* **19** (2006), 969–998.

[19] C. Topaz and A.L. Bertozzi, Swarming patterns in a two-dimensional kinematic model for biological groups, *SIAM Journal on Applied Mathematics* **65** (2004), 152–174.

[20] C.M. Topaz, A.L. Bertozzi, and M.A. Lewis, A nonlocal continuum model for biological aggregation, *Bulletin of Mathematical Biology* **68** (2006), 1601–1623.

[21] S. Ukai, On the existence of global solutions of mixed problem for non-linear Boltzmann equation, *Proceedings of the Japan Academy* **50** (1974), 179–184.

[22] S. Ukai and T. Yang, The Boltzmann equation in the space $L^2 \cap L^\infty_\beta$: Global and time-periodic solutions, *Analysis and Applications* **4** (2006), 263–310.

[23] T. Vicsek, A. Czirók, E. Ben-Jacob, and O. Shochet, Novel type of phase transition in a system of self-driven particles, *Phys. Rev. Lett.* **75** (1995), 1226–1229.

[24] C. Villani, A review of mathematical topics in collisional kinetic theory. *Handbook of mathematical fluid dynamics*, Vol. I, 71–305, North-Holland, Amsterdam, 2002.

[25] C. Villani, Hypocoercivity, *Memoirs Amer. Math. Soc.*, to appear.

[26] C. Yates, R. Erban, C. Escudero, L. Couzin, J. Buhl, L. Kevrekidis, P. Maini and D. Sumpter, Inherent noise can facilitate coherence in collective swarm motion, *Proceedings of the National Academy of Sciences* **106** (14) (2009), 5464–5469.