

A gradient flow approach to linear Boltzmann equations

GIADA BASILE, DARIO BENEDETTO AND LORENZO BERTINI

Abstract. We introduce a gradient flow formulation of linear Boltzmann equations. Under a diffusive scaling we derive a diffusion equation by using the machinery of gradient flows.

Mathematics Subject Classification (2010): 35Q20 (primary); 82C40, 49Q20 (secondary).

1. Introduction

The Boltzmann equation describes the evolution of the one-particle distribution on position and velocity of a rarefied gas. It has become a paradigmatic equation since it encodes most of the conceptual and technical issues in the description of the statistical properties for out of equilibrium systems. In particular, from a mathematical point of view, a global existence and uniqueness result is still lacking. In the kinetic regime, some transport phenomena can be described by linear Boltzmann equations. Typical examples are the charge (or mass) transport in the Lorentz gas [27], the evolution of a tagged particle in a Newtonian system in thermal equilibrium [35], and the propagation of lattice vibrations in insulating crystals [7]. Since the evolution equations are linear, their analysis is simpler. From one side, these equations have been derived from an underlying microscopic dynamics globally in time [7, 8, 13, 20]. From the other side, several results on the asymptotic behavior of the one-particle distribution have been obtained. In particular, by considering non degenerate scattering rates, under a diffusive rescaling the linear Boltzmann equation converges to the heat equation [3, 10, 19, 25].

In the present paper, inspired by the general theory in [2], we propose a formulation of linear Boltzmann equations in terms of gradient flows. Recently there have been some attempts to formulate the Fokker-Planck equation associated to continuous time reversible Markov chains, equivalently homogeneous linear kinetic

L. Bertini acknowledges the support by the PRIN 20155PAWZB “Large Scale Random Structures”.

Received November 14, 2018; accepted April 16, 2019.

Published online December 2020.

equations, as gradient flows [17, 28, 30], essentially in terms of energy variational inequalities, and the potential applications of this approach have yet to be fully investigated. The case of homogeneous non linear Boltzmann equations is considered in [18]. The present approach is based on an entropy dissipation inequality and can be applied naturally to the inhomogeneous case, that appears novel. In perspective, this approach could be adapted to the non linear, non homogeneous Boltzmann equation.

We consider a linear Boltzmann equation of the form

$$(\partial_t + b(v) \cdot \nabla_x) f(t, x, v) = \int \pi(dv') \sigma(v, v') [f(t, x, v') - f(t, x, v)], \quad (1.1)$$

where $\pi(dv)$ is a reference probability measure on the velocity space, b is the drift, $\sigma(v, v') \geq 0$ is the scattering kernel and f is the density of the one-particle distribution with respect to $dx \pi(dv)$. We assume the detailed balance condition, *i.e.* $\sigma(v, v') = \sigma(v', v)$. The entropy $\mathcal{H}(f) = \int dx \int \pi(dv) f \log f$ is a Lyapunov functional for the evolution (1.1), and the transport term do not affect its rate of decrease. This observation will allow to formulate (1.1) as the following entropy dissipation inequality

$$\mathcal{H}(f(T)) + \int_0^T dt \mathcal{E}(f(t)) + \mathcal{R}_0(f) \leq \mathcal{H}(f(0)), \quad (1.2)$$

where $\mathcal{E}(f) = \int dx \iint \pi(dv) \pi(dv') \sigma(v, v') [\sqrt{f(x, v')} - \sqrt{f(x, v)}]^2$ is the Dirichlet form of the square root of f and $\mathcal{R}_0 \geq 0$ is a kinematic term that will be defined later.

As an application of (1.2) we discuss the diffusive limit of the linear Boltzmann equation. This is a classical result, but the gradient flow formulation provides a transparent proof and allows to consider more general initial conditions, which are only required to satisfy an entropy bound. More precisely, we will show that in the diffusive scaling limit the particle density converges to the solution of the heat equation. The proof will be achieved by taking the limit in the rescaled entropy dissipation inequality and deducing the corresponding inequality for the heat equation.

ACKNOWLEDGEMENTS. We are grateful to Mauro Mariani for useful discussions about gradient flows and for his comments on an earlier version of the current manuscript.

2. A gradient flow formulation

In this section we introduce a gradient flow formulation of non-homogeneous linear kinetic equations. Both for ease of presentation and for future use, we however first

review the gradient flow formulation of the heat equation, that is here considered in somewhat different setting that includes the current as a dynamical variable.

Throughout the whole paper, the space domain is the d -dimensional torus $\mathbb{T}^d := \mathbb{R}^d / \mathbb{Z}^d$ and we denote by dx the Haar measure on \mathbb{T}^d . The set of Borel probability measures on \mathbb{T}^d is denoted by $\mathcal{P}(\mathbb{T}^d)$ that we consider endowed with the (metrizable) topology induced by the weak convergence. Recall finally that the *entropy* is the convex lower semicontinuous functional $H : \mathcal{P}(\mathbb{T}^d) \rightarrow [0, +\infty]$ defined by $H(\mu) = \int dx \rho \log \rho$ if $d\mu = \rho dx$ and $H(\mu) = +\infty$ otherwise.

Heat equation

We start by an informal discussion. Consider the heat equation on \mathbb{T}^d

$$\partial_t \rho = \nabla \cdot D \nabla \rho,$$

where ρ is a probability density and the diffusion coefficient D is a positive symmetric $d \times d$ matrix. We introduce the currents as vector fields on \mathbb{T}^d , denoted by j . Given ρ , we define the associated current $j^\rho := -D \nabla \rho$. We can then rewrite the heat equation as

$$\begin{cases} \partial_t \rho + \nabla \cdot j = 0 \\ j = j^\rho. \end{cases} \tag{2.1}$$

We shall rewrite this system as a variational inequality that expresses the decrease of the entropy.

Fix $T > 0$. On the set of paths $(\rho(t), j(t)), t \in [0, T]$, satisfying the continuity equation $\partial_t \rho + \nabla \cdot j = 0$ consider the action functional

$$I(\rho, j) = \frac{1}{2} \int_0^T dt \int dx \frac{1}{\rho(t)} [j(t) + D \nabla \rho(t)] \cdot D^{-1} [j(t) + D \nabla \rho(t)],$$

where \cdot denotes the inner product in \mathbb{R}^d . This functional arises naturally by analyzing the large deviation asymptotics of N independent Brownians [14, 23] and its connection with the gradient flow formulation of the heat equation is discussed in [1]. To be precise, the rate function in [1, 14, 23] does not include the current as a dynamical variable but it can be extended to this case, see [12] for a similar functional in the context of stochastic lattice gases.

Observe that $I \geq 0$ and $I(\rho, j) = 0$ if and only if $j = j^\rho$. Hence the second equation in (2.1) is equivalent to $I(\rho, j) \leq 0$. By expanding the square we deduce

$$\int_0^T dt \int dx \left[\frac{1}{2} \frac{1}{\rho(t)} j(t) \cdot D^{-1} j(t) + \frac{1}{2} \frac{1}{\rho(t)} \nabla \rho(t) \cdot D \nabla \rho(t) + \frac{1}{\rho(t)} \nabla \rho(t) \cdot j(t) \right] \leq 0. \tag{2.2}$$

Since $(\nabla \rho) / \rho = \nabla \log \rho$, integrating by parts and using the continuity equation, the last term is the total derivative of $H(\rho(t))$.

We now introduce the *Fisher information* E as the Dirichlet form of square root, namely

$$E(\rho) = \frac{1}{2} \int dx \frac{1}{\rho} \nabla \rho \cdot D \nabla \rho = 2 \int dx \nabla \sqrt{\rho} \cdot D \nabla \sqrt{\rho}.$$

Let also the *kinematic term* R be the functional on the set the path $(\rho(t), j(t))$ defined by

$$R(\rho, j) = \frac{1}{2} \int_0^T dt \int dx \frac{1}{\rho(t)} j(t) \cdot D^{-1} j(t),$$

then (2.2) reads

$$H(\rho(T)) + \int_0^T dt E(\rho(t)) + R(\rho, j) \leq H(\rho(0)), \tag{2.3}$$

which is the gradient flow formulation of the heat equation that we will use here.

We now specify the precise formulation in which we consider a family of probabilities $\mu_t(dx) = \rho(t, x) dx$, $t \in [0, T]$, while the currents are the vector valued measures $J(dt, dx) = j(t, x) dt dx$. Given $T > 0$ let $C([0, T]; \mathcal{P}(\mathbb{T}^d))$ be the set of continuous paths on $\mathcal{P}(\mathbb{T}^d)$ endowed with the topology of uniform convergence. Let also $\mathcal{M}([0, T] \times \mathbb{T}^d; \mathbb{R}^d)$ be the set of vector valued Radon measures on $[0, T] \times \mathbb{T}^d$ endowed with the weak* topology. Set $S := C([0, T]; \mathcal{P}(\mathbb{T}^d)) \times \mathcal{M}([0, T] \times \mathbb{T}^d; \mathbb{R}^d)$ endowed with the product topology.

Given a positive $d \times d$ matrix D , the Fisher information $E : \mathcal{P}(\mathbb{T}^d) \rightarrow [0, \infty]$ can be defined by the variational formula

$$E(\mu) = 2 \sup_{\phi \in C^2(\mathbb{T}^d)} \left\{ - \int d\mu e^{-\phi} \nabla \cdot D \nabla e^{\phi} \right\}, \tag{2.4}$$

which implies its lower semicontinuity and convexity.

The kinematic term $R : S \rightarrow [0, \infty]$ admits the variational representation

$$R(\mu, J) = \sup_{w \in C([0, T] \times \mathbb{T}^d; \mathbb{R}^d)} \left\{ J(w) - \frac{1}{2} \int_0^T dt \int d\mu_t w \cdot Dw \right\}, \tag{2.5}$$

which implies its lower semicontinuity and convexity.

Definition 2.1. Let $\nu \in \mathcal{P}(\mathbb{T}^d)$ with $H(\nu) < +\infty$. A path $(\mu, J) \in S$ is a solution of the heat equation with initial condition ν iff $\mu_0 = \nu$ and

$$\int_0^T dt \mu_t(\partial_t \phi) + J(\nabla \phi) = 0, \quad \phi \in C_c^1((0, T) \times \mathbb{T}^d) \tag{2.6}$$

$$H(\mu_T) + \int_0^T dt E(\mu_t) + R(\mu, J) \leq H(\nu). \tag{2.7}$$

The standard formulation of the heat equation as gradient flow of the entropy is recovered from (2.7) by projecting on the density. Indeed, by the Benamou-Brenier lemma [9], we deduce that if (μ, J) is a solution to the heat equation according to Definition 2.1, then $\mu = (\mu_t)_{t \in [0, T]}$ satisfies

$$H(\mu_T) + \int_0^T dt \left\{ E(\mu_t) + \frac{1}{2} |\dot{\mu}_t|^2 \right\} \leq H(v), \tag{2.8}$$

where $|\dot{\mu}_t|$ is the metric derivative of $t \mapsto \mu_t$ with respect to the Wasserstein-2 distance, namely $|\dot{\mu}_t| = \lim_{h \rightarrow 0} d_{W_2}(\mu_{t+h}, \mu_t) / h$, where d_{W_2} denotes the Wasserstein-2 distance on $\mathcal{P}(\mathbb{T}^d)$.

Conversely, let μ be a solution to (2.8) satisfying $\mu_0 = v$. Introduce the functional J^μ on $C^1([0, T] \times \mathbb{T}^d; \mathbb{R}^d)$ defined by $J^\mu(w) = \int_0^T dt \mu_t(\nabla \cdot Dw_t)$. Since $\int_0^T dt E(\mu_t) \leq H(v)$, the functional J^μ extends to an element of $\mathcal{M}([0, T] \times \mathbb{T}^d; \mathbb{R}^d)$, still denoted by J^μ . Using again the Benamou-Brenier lemma it is then straightforward to check that the pair (μ, J^μ) is a solution to the heat equation in the sense of Definition 2.1.

The previous remarks, together with the existence and uniqueness result for the formulation (2.8) in [21], imply the following statement.

Proposition 2.2. *For each $v \in \mathcal{P}(\mathbb{T}^d)$, with $H(v) < \infty$, there exists a unique solution of the heat equation with initial condition v .*

Linear Boltzmann equations

We do not need any particular hypotheses on the velocity space \mathcal{V} that is assumed to be a Polish space, *i.e.* a metrizable complete and separable topological space. We denote by $\mathcal{P}(\mathbb{T}^d \times \mathcal{V})$ the set of probabilities on $\mathbb{T}^d \times \mathcal{V}$, that we consider endowed with the topology of weak convergence. We suppose given a Borel probability measure π on \mathcal{V} , a symmetric scattering kernel σ , *i.e.* a Borel function $\sigma : \mathcal{V} \times \mathcal{V} \rightarrow [0, +\infty)$ satisfying $\sigma(v, v') = \sigma(v', v)$, $v, v' \in \mathcal{V}$, and a drift $b : \mathcal{V} \rightarrow \mathbb{R}^d$. Given $P \in \mathcal{P}(\mathbb{T}^d \times \mathcal{V})$, we denote by $\mathcal{H}(P)$ the relative entropy of P with respect to the probability $dx \pi(dv)$ namely, $\mathcal{H}(P) = \int \int dx \pi(dv) f \log f$ if $dP = f dx \pi(dv)$ and $\mathcal{H}(P) = +\infty$ otherwise.

Also in this case we start by an informal discussion. Fix $T > 0$. Given a path $(P_t)_{t \in [0, T]}$ on $\mathcal{P}(\mathbb{T}^d \times \mathcal{V})$ with $dP_t = f(t, x, v) dx \pi(dv)$, we use the shorthand notation $f = f(t, x, v)$, $f' = f(t, x, v')$ and set

$$\eta^f = \eta^f(t, x, v, v') := \sigma(f - f') = \sigma(v, v')[f(t, x, v) - f(t, x, v')]. \tag{2.9}$$

We then rewrite the linear Boltzmann equation (1.1) in the form

$$\begin{cases} (\partial_t + b(v) \cdot \nabla_x) f(t, x, v) + \int \pi(dv') \eta(t, x, v, v') = 0 \\ \eta = \eta^f \end{cases} \tag{2.10}$$

We understand that the first equation has to be satisfied weakly and we shall refer to it as the *balance* equation. We are going to rewrite the condition $\eta = \eta^f$ as an inequality that expresses the decrease of the relative entropy \mathcal{H} . To this end, given $\varkappa \geq 0$ let $\Phi_\varkappa: \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \rightarrow [0, +\infty)$ be the convex function defined by

$$\Phi_\varkappa(p, q; \xi) := \sup_{\lambda \in \mathbb{R}} \left\{ \lambda \xi - \varkappa p (e^\lambda - 1) - \varkappa q (e^{-\lambda} - 1) \right\},$$

observing that given $p, q \in \mathbb{R}_+$ the map $\xi \mapsto \Phi_\varkappa(p, q; \xi)$ is positive (take $\lambda = 0$), and equal to zero iff $\xi = \varkappa(p - q)$. Explicitly, as few computations show, Φ_\varkappa reads

$$\begin{aligned} \Phi_\varkappa(p, q; \xi) = \xi \left[\operatorname{ash} \frac{\xi}{2\varkappa\sqrt{pq}} - \operatorname{ash} \frac{\varkappa(p - q)}{2\varkappa\sqrt{pq}} \right] \\ - \left[\sqrt{\xi^2 + 4\varkappa^2 pq} - \sqrt{[\varkappa(p - q)]^2 + 4\varkappa^2 pq} \right], \end{aligned} \tag{2.11}$$

where we recall that $\operatorname{ash}(z) = \log(z + \sqrt{1 + z^2})$. We note that if $\varkappa = 0$ then $\Phi_0(p, q; 0) = 0$ while $\Phi_0(p, q; \xi) = +\infty$ if $\xi \neq 0$.

Fix a path $(f(t), \eta(t)), t \in [0, T]$ satisfying $\eta(t, x, v, v') = -\eta(t, x, v', v)$ and the balance equation in (2.10). The condition $\eta(t) = \eta^{f(t)}, t \in [0, T]$ is equivalent to

$$\mathcal{I}(f, \eta) := \int_0^T dt \int dx \iint \pi(dv) \pi(dv') \Phi_\sigma(f, f'; \eta) \leq 0. \tag{2.12}$$

This functional is connected with the large deviations asymptotic of a Markov chain on \mathcal{V} with transition rates $\sigma(v', v)\pi(dv')$, see [11, 31].

We next write

$$\Phi_\varkappa(p, q; \xi) = \Phi_\varkappa(p, q; 0) + \xi \frac{\partial}{\partial \xi} \Phi_\varkappa(p, q; 0) + \Psi_\varkappa(p, q; \xi). \tag{2.13}$$

By a few explicit computations,

$$\begin{aligned} \Phi_\varkappa(p, q; 0) &= \varkappa(\sqrt{p} - \sqrt{q})^2 \\ \frac{\partial}{\partial \xi} \Phi_\varkappa(p, q; 0) &= \frac{1}{2} \log \frac{q}{p} \\ \Psi_\varkappa(p, q; \xi) &= \xi \operatorname{ash} \frac{\xi}{2\varkappa\sqrt{pq}} - \left[\sqrt{\xi^2 + 4\varkappa^2 pq} - 2\varkappa\sqrt{pq} \right]. \end{aligned}$$

Observe that Ψ_\varkappa has the variational representation

$$\Psi_\varkappa(p, q; \xi) = \sup_{\lambda \in \mathbb{R}} \left\{ \lambda \xi - 2\varkappa\sqrt{pq} [\operatorname{ch} \lambda - 1] \right\}. \tag{2.14}$$

In particular, $\Psi_\varkappa \geq 0$. Moreover, while the map $(p, q; \xi) \mapsto \Psi_\varkappa(p, q; \xi)$ is convex, the map $\xi \mapsto \Psi_\varkappa(p, q; \xi)$ is strictly convex. Finally, $\Psi_\varkappa(p, q; \xi) \sim \xi^2$ for ξ small and $\Psi_\varkappa(p, q; \xi) \sim |\xi| \log |\xi|$ for ξ large.

Observe now that for the path $(f(t), \eta(t)), t \in [0, T]$ satisfying the balance equation in (2.10) we have

$$\begin{aligned} \frac{d}{dt} \mathcal{H}(f(t)) &= \int dx \int \pi(dv) \log f \left[-b(v) \cdot \nabla_x f - \int \pi(dv') \eta(t, x, v, v') \right] \\ &= - \int dx \iint \pi(dv) \pi(dv') \eta \log f, \end{aligned}$$

since the first term is a total derivative in x . Hence, by the antisymmetry of η ,

$$\int dx \iint \pi(dv) \pi(dv') \eta \frac{\partial}{\partial \xi} \Phi_\varkappa(f, f'; 0) = \frac{d}{dt} \mathcal{H}(f(t)).$$

for any $\varkappa > 0$. Setting $\varkappa = \sigma$, inserting (2.13) and integrating in time we obtain that, for any $(f(t), \eta(t)), t \in [0, T]$ satisfying the balance equation, it holds

$$\begin{aligned} \mathcal{H}(f(T)) + \int_0^T dt \int dx \iint \pi(dv) \pi(dv') [\Phi_\sigma(f, f'; 0) + \Psi_\sigma(f, f'; \eta)] \\ = \mathcal{H}(f(0)) + \int_0^T dt \int dx \iint \pi(dv) \pi(dv') \Phi_\sigma(f, f'; \eta). \end{aligned} \tag{2.15}$$

Gathering the above computations we conclude that (2.12) can be rewritten as

$$\mathcal{H}(f(T)) + \int_0^T dt \mathcal{E}(f(t)) + \mathcal{R}(f, \eta) \leq \mathcal{H}(f(0)), \tag{2.16}$$

where

$$\mathcal{E}(f) = \int dx \iint \pi(dv) \pi(dv') \sigma(v, v') [\sqrt{f'} - \sqrt{f}]^2 \tag{2.17}$$

and

$$\mathcal{R}(f, \eta) = \int_0^T dt \int dx \iint \pi(dv) \pi(dv') \Psi_\sigma(f, f'; \eta). \tag{2.18}$$

The inequality (2.16), formally analogous to (2.3), is the proposed gradient flow formulation of the linear Boltzmann equation (1.1).

We now discuss the precise formulation in which we introduce the measures $dP = f(x, v) dx \pi(dv)$ and $\Theta(dt, dx, dv, dv') = \eta(t, x, v, v') dt dx dv dv'$. We first specify the hypotheses on the scattering rate σ and the drift b that are assumed to hold throughout the whole paper.

Assumption 2.3.

- (i) The scattering kernel is a Borel function $\sigma : \mathcal{V} \times \mathcal{V} \rightarrow [0, +\infty)$ satisfying $\sigma(v, v') = \sigma(v', v), (v, v') \in \mathcal{V} \times \mathcal{V}$.

- (ii) The *scattering rate* $\lambda: \mathcal{V} \rightarrow [0, +\infty)$ is defined by $\lambda(v) := \int \pi(dv') \sigma(v, v')$. We require that it has all exponential moments with respect to π namely, $\pi[e^{\gamma\lambda}] < +\infty$ for any $\gamma \in \mathbb{R}_+$.
- (iii) The *drift* is a Borel function $b: \mathcal{V} \rightarrow \mathbb{R}^d$. We require that it has all exponential moments with respect to π namely, $\pi[e^{\gamma|b|}] < +\infty$ for any $\gamma \in \mathbb{R}_+$, where $|b|$ is the Euclidean norm of b .

Given $T > 0$ let $C([0, T]; \mathcal{P}(\mathbb{T}^d \times \mathcal{V}))$ be the set of continuous paths on $\mathcal{P}(\mathbb{T}^d \times \mathcal{V})$ endowed with the topology of uniform convergence. Denote by $\mathcal{M}_a([0, T] \times \mathbb{T}^d \times \mathcal{V} \times \mathcal{V})$ the set of finite Radon measures on $[0, T] \times \mathbb{T}^d \times \mathcal{V} \times \mathcal{V}$ antisymmetric with respect to the exchange of the last two variables endowed with the weak* topology. Set $\mathcal{S} := C([0, T]; \mathcal{P}(\mathbb{T}^d \times \mathcal{V})) \times \mathcal{M}_a([0, T] \times \mathbb{T}^d \times \mathcal{V} \times \mathcal{V})$ endowed with the product topology. Let also $C_{be}([0, T]; \mathcal{P}(\mathbb{T}^d \times \mathcal{V}))$ the set of paths $(P_t)_{t \in [0, T]}$ in $C([0, T]; \mathcal{P}(\mathbb{T}^d \times \mathcal{V}))$ such that $\sup_{t \in [0, T]} \mathcal{H}(P_t) < +\infty$ and let finally $\mathcal{S}_{be} := C_{be}([0, T]; \mathcal{P}(\mathbb{T}^d \times \mathcal{V})) \times \mathcal{M}_a([0, T] \times \mathbb{T}^d \times \mathcal{V} \times \mathcal{V})$

If $P \in \mathcal{P}(\mathbb{T}^d \times \mathcal{V})$ has finite entropy, the Dirichlet form of the square root \mathcal{E} can be defined by the variational formula

$$\mathcal{E}(P) := 2 \sup_{\phi \in C_b(\mathbb{T}^d \times \mathcal{V})} \iint P(dx, dv) \pi(dv') \sigma(v, v') [1 - e^{\phi(x, v') - \phi(x, v)}]. \tag{2.19}$$

Note indeed the right-hand side is well defined for any $\phi \in C_b(\mathbb{T}^d \times \mathcal{V})$ in view of Assumption 2.3 and the basic entropy inequality $P(\psi) \leq \mathcal{H}(P) + \log \int dx \pi(dv) e^\psi$, $\psi: \mathbb{T}^d \times \mathcal{V} \rightarrow \mathbb{R}$. The representation (2.19) corresponds to the Donsker-Varadhan large deviation for the empirical measure of the continuous time Markov chain on \mathcal{V} with transition rates $\sigma(v', v)\pi(dv')$ [16]. Indeed, $\mathcal{E}(P) = \sup_\phi \{-P(e^{-\phi} \mathcal{L} e^\phi)\}$, where

$$\mathcal{L}g(v) = \int \pi(dv') \sigma(v', v) [g(v') - g(v)]. \tag{2.20}$$

A variational representation for the kinematic term \mathcal{R} is obtained by combining (2.14) with the simple observation that for $p, q \in \mathbb{R}_+$ we have $-2\sqrt{pq} = \sup_{a>0} \{-ap - a^{-1}q\}$. We thus let $\mathcal{R}: \mathcal{S}_{be} \rightarrow [0, +\infty]$ be the functional defined by

$$\begin{aligned} \mathcal{R}(P, \Theta) := & \sup_{\zeta, \alpha} \left\{ \Theta(\zeta) - \int_0^T dt \iiint P_t(dx, dv) \pi(dv') \sigma(v, v') \right. \\ & \left. \times [\text{ch } \zeta(t, x, v, v') - 1][\alpha(t, x, v, v') + \alpha(t, x, v', v)^{-1}] \right\}, \end{aligned} \tag{2.21}$$

where the supremum is carried out over the continuous functions $\zeta: [0, T] \times \mathbb{T}^d \times \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ with compact support and antisymmetric with respect to the exchange of the last two variables and the bounded continuous functions $\alpha: [0, T] \times \mathbb{T}^d \times \mathcal{V} \times$

$\mathcal{V} \rightarrow (0, +\infty)$ uniformly bounded away from zero. As before, the basic entropy inequality implies that \mathcal{R} is well defined.

At this point the gradient flow formulation of the linear Boltzmann equations is simply specified by the following entropy dissipation inequality.

Definition 2.4. Let $Q \in \mathcal{P}(\mathbb{T}^d \times \mathcal{V})$ with $\mathcal{H}(Q) < +\infty$. An element $(P, \Theta) \in \mathcal{S}_{\text{be}}$ is a solution to the linear Boltzmann equation with initial condition Q if and only if $P_0 = Q$ and

$$\int_0^T dt P_t(\partial_t \phi + b \cdot \nabla_x \phi) = \frac{1}{2} \int \Theta(dt, dx, dv, dv') [\phi(t, x, v) - \phi(t, x, v')], \tag{2.22}$$

$$\mathcal{H}(P_T) + \int_0^T dt \mathcal{E}(P_t) + \mathcal{R}(P, \Theta) \leq \mathcal{H}(Q). \tag{2.23}$$

for all continuous functions $\phi: (0, T) \times \mathbb{T}^d \times \mathcal{V}$ with compact support and continuously differentiable in the first two variables.

Remark 2.5. If (P, Θ) is a solution to the linear Boltzmann equation in the time interval $[0, T]$ then it solves the same problem in the time interval $[0, t]$, $t \leq T$ as well. This follows from the fact that any element $(P, \Theta) \in \mathcal{S}_{\text{be}}$ satisfies, for $0 \leq s < t \leq T$, the inequality

$$\mathcal{H}(P_t) + \int_s^t du \mathcal{E}(P_u) + \mathcal{R}^{s,t}(P, \Theta_{[s,t]}) \geq \mathcal{H}(P_s), \tag{2.24}$$

where $\Theta_{[s,t]}$ is the restriction of Θ to the interval $[s, t]$ and the kinematic term $\mathcal{R}^{s,t}$ is defined as in (2.21) with the interval $[0, T]$ replaced by $[s, t]$. This inequality corresponds in fact to the trivial inequality $\mathcal{I}_{[s,t]}(P, \Theta) \geq 0$ where the action functional $\mathcal{I}_{[s,t]}$ is defined as in (2.12) with the interval $[0, T]$ replaced by $[s, t]$. The actual proof of (2.24) is detailed in Appendix A.

It is of course possible to obtain a formulation only in terms of the one particle distribution. More precisely, the formulation (1.2) is obtained from (2.23) simply by letting $\mathcal{R}_0: C_{\text{be}}([0, T]; \mathcal{P}(\mathbb{T}^d \times \mathcal{V})) \rightarrow [0, +\infty]$ be the functional defined by $\mathcal{R}_0(P) = \inf_{\Theta} \mathcal{R}(P, \Theta)$ where the infimum is carried out over all $\Theta \in \mathcal{M}_a([0, T] \times \mathbb{T}^d \times \mathcal{V} \times \mathcal{V})$ such that the pair (P, Θ) satisfies the balance equation (2.22). It is however unclear to us whether $\mathcal{R}_0(P)$ could be represented by the metric derivative of $t \mapsto P_t$ with respect to a suitable distance on $\mathcal{P}(\mathbb{T}^d \times \mathcal{V})$.

We now state an existence and uniqueness result for the above formulation, together with a continuous dependence on the initial condition and the coefficients. In particular uniqueness implies that solutions to (1.1) with bounded entropy are characterized by the gradient flow formulation in Definition 2.4.

Theorem 2.6. *For each $Q \in \mathcal{P}(\mathbb{T}^d \times \mathcal{V})$, with $\mathcal{H}(Q) < +\infty$, there exists a unique solution (P, Θ) to the linear Boltzmann equation with initial condition Q . Furthermore:*

- (i) Set $\Theta^P(dt, dx, dv, dv') := dt \sigma(v, v') [P_t(dx, dv)\pi(dv') - \pi(dv)P_t(dx, dv')]$. Then $\Theta = \Theta^P$;
- (ii) Let $\{Q^n\} \subset \mathcal{P}(\mathbb{T}^d \times \mathcal{V})$ be such that $Q^n \rightarrow Q$ and $\mathcal{H}(Q^n) \rightarrow \mathcal{H}(Q)$ and denote by $(P^n, \Theta^n) \in \mathcal{S}_{be}$ the solution to the linear Boltzmann equation with initial condition Q^n . Then the sequence $\{(P^n, \Theta^n)\}$ converges to (P, Θ) ;
- (iii) Fix $Q \in \mathcal{P}(\mathbb{T}^d \times \mathcal{V})$ with $\mathcal{H}(Q) < +\infty$ and consider coefficients b, σ together with sequences b^n, σ^n all satisfying Assumption 2.3. Denote by $(P^n, \Theta^n) \in \mathcal{S}_{be}$ the solution to the linear Boltzmann equation with initial condition Q and coefficients b^n, σ^n . If $b_n \rightarrow b$ in π probability, $\sigma^n \rightarrow \sigma$ in $\pi \times \pi$ probability, and $\lim_n \{ \log \pi [e^{\gamma|b^n - b|}] + \log \pi [e^{\gamma|\lambda^n - \lambda|}] \} = 0$ for any $\gamma > 0$ then the sequence $\{(P^n, \Theta^n)\}$ converges to (P, Θ) .

While uniqueness will be proven using the argument in [21], the key ingredient for the continuity result is the following lemma. Its proof, whose details are omitted, is achieved by truncating with continuous and bounded functions and using the basic entropy inequality.

Lemma 2.7. *Let $\{P^n\} \subset \mathcal{P}(\mathbb{T}^d \times \mathcal{V})$ be a sequence converging to P and satisfying the entropy bound $\sup_n \mathcal{H}(P^n) < +\infty$. Then $P^n(\phi) \rightarrow P(\phi)$ for any function ϕ having all exponential moments with respect to π . Moreover, if ϕ has all exponential moments and $\lim_n \log \pi [e^{\gamma|\phi^n - \phi|}] = 0$ for any $\gamma > 0$ then $P^n(\phi^n) \rightarrow P(\phi)$.*

In view of the variational definition (2.19), this lemma readily implies that the Dirichlet form \mathcal{E} is lower semicontinuous on sublevel sets of the entropy. Analogously, recalling (2.21), the kinematic term \mathcal{R} is lower semicontinuous on the sets $\{(P, \Theta) \in \mathcal{S}_{be} : \sup_{t \in [0, T]} \mathcal{H}(P_t) \leq \ell\}$, $\ell \in \mathbb{R}_+$.

Proof of Theorem 2.6. We start by proving uniqueness, and in particular by showing that if (P^1, Θ^1) and (P^2, Θ^2) are solutions then $P^1 = P^2$. Assume by contradiction that there exists $t \in (0, T]$ such that $P_t^1 \neq P_t^2$ and let $(\bar{P}, \bar{\Theta}) = \frac{1}{2}(P^1, \Theta^1) + \frac{1}{2}(P^2, \Theta^2)$. In view of Remark 2.5, by the convexity of \mathcal{E} , \mathcal{R} and the strict convexity of \mathcal{H}

$$\mathcal{H}(\bar{P}_t) + \int_0^t ds \mathcal{E}(\bar{P}_s) + \mathcal{R}^{0,t}(\bar{P}, \bar{\Theta}) < \mathcal{H}(Q),$$

which by (2.24) provides the desired contradiction. Uniqueness is now concluded by observing that, for a given $P \in C_{be}([0, T]; \mathcal{P}(\mathbb{T}^d \times \mathcal{V}))$, the map $\Theta \mapsto \mathcal{R}(P, \Theta)$ is strictly convex, so that we can repeat the argument above with $P^1 = P^2 = P$ and deduce $\Theta^1 = \Theta^2$.

Postponing the proof of the existence, we show item (i). We write $\Theta = \Theta^P + \tilde{\Theta}$ and we observe that, in view of the balance equation (2.22), $\tilde{\Theta}(\zeta) = 0$ if $\zeta(t, x, v, v') = z(t, x, v') - z(t, x, v)$ for some function z . By choosing in the variational formula (2.21) $\zeta = z' - z$, where $z = z(t, x, v)$ and $z' = z(t, x, v')$, we

get

$$\begin{aligned} \mathcal{R}(P, \Theta^P + \tilde{\Theta}) &\geq \sup_{z, \alpha} \left\{ \Theta^P(z' - z) - \int_0^T dt \iiint P_t(dx, dv) \pi(dv') \sigma(v, v') \right. \\ &\quad \left. \times [\text{ch}(z' - z) - 1][\alpha(t, x, v, v') + \alpha(t, x, v', v)^{-1}] \right\} \\ &= \mathcal{R}(P, \Theta^P), \end{aligned}$$

where the last equality follows by direct computation. We conclude by uniqueness.

We next prove item (ii). Remark 2.5 implies that

$$\overline{\lim}_{n \rightarrow \infty} \sup_{t \in [0, T]} \mathcal{H}(P_t^n) \leq \mathcal{H}(Q). \tag{2.25}$$

In view of the lower semicontinuity of \mathcal{H} and the observation after Lemma 2.7 regarding \mathcal{E} and \mathcal{R} , using the uniqueness it is enough to show precompactness of the sequence $\{(P^n, \Theta^n)\} \subset \mathcal{S}_{\text{be}}$. Observe indeed that, by Assumption 2.3 and Lemma 2.7, we can take the limit $n \rightarrow \infty$ in the balance equation (2.22).

To prove precompactness of $\{\Theta^n\}$, observe that from (2.23) and the variational representation (2.21) it follows

$$\sup_n \sup_{\zeta: \|\zeta\|_\infty \leq 1} \Theta^n(\zeta) < +\infty$$

and we conclude by the Banach-Alaoglu theorem.

The bound (2.25) implies, by the coercive properties of the relative entropy and Prohorov theorem, that there exist a compact $\mathcal{K} \subset \subset \mathcal{P}(\mathbb{T}^d \times \mathcal{V})$ such that $P_t^n \in \mathcal{K}$ for any n and $t \in [0, T]$. Hence, by Ascoli-Arzelà theorem, to prove the precompactness of $\{P^n\}$ it is enough to show that for each continuous $g: \mathbb{T}^d \times \mathcal{V} \rightarrow \mathbb{R}$ with compact support and continuously differentiable with respect to x we have

$$\lim_{\delta \downarrow 0} \sup_n \sup_{|t-s| < \delta} |P_t^n(g) - P_s^n(g)| = 0. \tag{2.26}$$

From the balance equation (2.22) we deduce

$$\begin{aligned} P_t^n(g) - P_s^n(g) &= - \int_s^t d\tau P_\tau^n(b \cdot \nabla_x g) \\ &\quad - \frac{1}{2} \int_{[s, t] \times \mathbb{T}^d \times \mathcal{V} \times \mathcal{V}} \Theta^n(d\tau, dx, dv, dv') [g(x, v) - g(x, v')]. \end{aligned}$$

By Assumption 2.3, (2.25), and the basic entropy inequality, the first term on the right-hand side vanishes as $|t - s| \rightarrow 0$ uniformly in n . On the other hand, by choosing $\alpha = 1$ in the variational representation (2.21),

$$\begin{aligned} &\left| \int_{[s, t] \times \mathbb{T}^d \times \mathcal{V} \times \mathcal{V}} \Theta^n(d\tau, dx, dv, dv') [g(x, v) - g(x, v')] \right| \leq \mathcal{R}(P^n, \Theta^n) \\ &+ 2 \int_s^t d\tau \iiint P_\tau(dx, dv) \pi(dv') \sigma(v, v') \{ \text{ch}[g(x, v) - g(x, v')] - 1 \}. \end{aligned}$$

Replacing g by γg with $\gamma > 0$, using (2.25), Assumption 2.3, the basic entropy inequality, and $\mathcal{R}(P^n, \Theta^n) \leq \mathcal{H}(Q^n)$, we obtain that there exists a constant C independent on n, t, s such that

$$\begin{aligned} & \left| \int_{[s,t] \times \mathbb{T}^d \times \mathcal{V} \times \mathcal{V}} \Theta^n(d\tau, dx, dv, dv') [g(x, v) - g(x, v')] \right| \\ & \leq \frac{1}{\gamma} \sup_n \mathcal{H}(Q^n) + \frac{C}{\gamma} |t - s| \exp\{2\gamma \|g\|_\infty\}. \end{aligned}$$

By choosing $\gamma = (2\|g\|_\infty)^{-1} \log(1/|t - s|)$ when $|t - s| \leq 1$ the bound (2.26) follows.

In view of the second statement in Lemma 2.7, the proof of item (iii) is achieved by the same arguments.

We finally prove the existence result. Consider first the case in which b and σ are continuous and bounded. If $Q(dx, dv) = f_0(x, v) dx \pi(dv)$ for some continuous f_0 uniformly bounded away from zero, by classical results, the linear Boltzmann equation (1.1) has a continuous solution $f(t, x, v)$ uniformly bounded away from zero. Set $P_t(dx, dv) := f(t, x, v) dx \pi(dv)$ and $\Theta(dt, dx, dv, dv') := dt dx \pi(dv) \pi(dv') \sigma(v, v') [f(t, x, v) - f(t, x, v')]$. It is then straightforward to justify the informal computations presented before and deduce that (P, Θ) solves the linear Boltzmann equation according to Definition 2.4. Existence in the general case of Q satisfying the relative entropy bound $\mathcal{H}(Q) < +\infty$ and for coefficients b and σ satisfying Assumption 2.3 is then achieved by items (ii) and (iii). \square

3. Diffusive limit

In this section we discuss the asymptotic behavior of linear Boltzmann equation, showing that in the diffusive scaling limit the marginal distribution of the position evolves according to the heat flow. This is a classical topic and has been much investigated in the literature, see [3, 10, 19, 25] and [15] for a more general setting. We observe that this issue has natural counterpart in probabilistic terms namely, the central limit for additive functional of Markov chains. Indeed, the linear Boltzmann equation (1.1) is the Fokker-Planck equation for the Markov process (V_t, X_t) where V_t is the continuous time Markov chain on \mathcal{V} with transition rates $\sigma(v', v) \pi(dv')$ while X_t is the \mathbb{R}^d -valued additive functional $X_t = \int_0^t ds b(V_s)$. We refer to [24] for a recent monograph on this topic.

The gradient flow formulation of linear Boltzmann equations discussed before allows a novel approach to the analysis of the diffusive limit. According to a general scheme formalized in [33, 34], a gradient flow formulation is particularly handy for analyzing asymptotic evolutions and it does not require a direct analysis of the dynamics. Indeed, by comparing Definition 2.4 and Definition 2.1 we realize that the balance equation (2.22) immediately leads to the continuity equation (2.6). Moreover, taking into account the convexity and lower semicontinuity of the entropy, in order to establish the diffusive limit we only need to prove two limiting

variational inequalities comparing the Dirichlet form and the kinematic term for the linear Boltzmann equation with the corresponding ones for the heat flow. We shall prove these variational inequalities but, maybe surprisingly, the two terms exchange their role in the diffusive limit: the Dirichlet form \mathcal{E} leads to R while the kinematic term \mathcal{R} leads to the Fisher information E .

To carry out the analysis of the diffusive limit of linear Boltzmann equations a few extra conditions, implying in particular homogenization of the velocity, are needed. As we show in the next section, this assumptions are satisfied for few natural models. To this end, recalling that the scattering rate λ is defined by $\lambda(v) = \int \pi(dv')\sigma(v', v)$, let $\tilde{\pi}$ be the probability on \mathcal{V} defined by

$$\tilde{\pi}(dv) := \frac{\lambda(v)}{\pi(\lambda)}\pi(dv). \tag{3.1}$$

Assumption 3.1.

- (i) The drift $b : \mathcal{V} \rightarrow \mathbb{R}^d$ is centered with respect to the measure π , namely $\pi(b) = 0$.
- (ii) The scattering rate λ satisfies $\pi[\lambda = 0] = 0$.
- (iii) $|b|^2/\lambda$ has all exponential moments, *i.e.* $\pi[\exp\{\gamma|b|^2/\lambda\}] < +\infty$ for any $\gamma > 0$.
- (iv) There exists a constant $C_0 > 0$ such that for any $g \in L^2(\tilde{\pi})$

$$\int d\tilde{\pi}[g - \tilde{\pi}(g)]^2 \leq C_0 \iint \tilde{\pi}(dv)\tilde{\pi}(dv') \frac{\sigma(v, v')}{\lambda(v)\lambda(v')} [g(v) - g(v')]^2. \tag{3.2}$$

We remark that, as in the case of phonon Boltzmann equation, in item (ii) we allow the case in which $\lambda(v) = 0$ for some $v \in \mathcal{V}$. Item (iv) corresponds to the assumption that the continuous time Markov chain with transition rates $\frac{\sigma(v, v')}{\lambda(v)\lambda(v')}\tilde{\pi}(dv')$ has spectral gap. The generator of this Markov chain is $(K - \mathbb{I})$, where K is given by

$$(Kg)(v) = \pi(\lambda) \int \tilde{\pi}(dv') \frac{\sigma(v, v')}{\lambda(v)\lambda(v')} g(v'). \tag{3.3}$$

Observe that

$$-(\mathcal{L}f)(v) = \lambda(v)[(\mathbb{I} - K)f](v), \tag{3.4}$$

where \mathcal{L} is the generator of the original Markov chain as defined in (2.20). We emphasize that we do not assume the spectral gap of the generator \mathcal{L} , in fact the linear phonon Boltzmann equation, that will be discussed in the next section, meets the requirements in Assumption 3.1 but its generator has no spectral gap.

Assumption 3.1 implies that there exists $\xi \in L^2(\tilde{\pi}; \mathbb{R}^d)$ such that $-\mathcal{L}\xi = b$. Indeed, item (i) implies that b/λ is centered with respect to $\tilde{\pi}$, item (ii) implies that $b/\lambda \in L^2(\tilde{\pi}; \mathbb{R}^d)$, and finally item (iii) implies that $\xi := (\mathbb{I} - K)^{-1}(b/\lambda) \in L^2(\tilde{\pi}; \mathbb{R}^d)$.

We need another technical condition on which we will rely to carry out a truncation on ξ .

Assumption 3.2. One of the following alternatives holds:

- (i) $(-\mathcal{L})^{-1}b$ is bounded; or
- (ii) there exists $C < \infty$ such that $\|(\mathbb{I} - K)^{-1}f\|_\infty \leq C\|f\|_\infty$, for any f such that $\tilde{\pi}(f) = 0$.

Observe that if the map $(v, v') \mapsto \frac{\sigma(v, v')}{\lambda(v)\lambda(v')}$ is continuous and bounded then alternative (ii) holds.

For notational convenience, in this section we formulate the linear Boltzmann equation in terms of the density (f, η) , where $dP_t = f(t) dx \pi(dv)$ and $d\Theta = \eta dt dx \pi(dv)\pi(dv')$. Indeed, the boundedness of the entropy implies the existence of f , while the boundedness of the metric term \mathcal{R} implies the existence of η .

Let $\varepsilon > 0$ be a scaling parameter and consider the linear Boltzmann equation on a torus of linear size ε^{-1} , equipped with the uniform probability distribution, on the time interval $[0, \varepsilon^{-2}T]$. Under a diffusive rescaling of space and time, the rescaled solution $(f^\varepsilon, \eta^\varepsilon)$ is defined on the torus of linear size one on the time interval $[0, T]$. It solves

$$\partial_t f^\varepsilon(t, x, v) + \frac{1}{\varepsilon} b(v) \cdot \nabla_x f^\varepsilon(t, x, v) + \frac{1}{\varepsilon^2} \int \pi(dv') \eta^\varepsilon(t, x, v, v') = 0 \tag{3.5}$$

$$\mathcal{H}(f^\varepsilon(T)) + \frac{1}{\varepsilon^2} \int_0^T dt \mathcal{E}(f^\varepsilon(t)) + \frac{1}{\varepsilon^2} \mathcal{R}(f^\varepsilon, \eta^\varepsilon) \leq \mathcal{H}(f_0^\varepsilon). \tag{3.6}$$

We set

$$\begin{aligned} \rho^\varepsilon(t, x) &:= \int \pi(dv) f^\varepsilon(t, x, v) \\ j^\varepsilon(t, x) &:= \frac{1}{\varepsilon} \int \pi(dv) f^\varepsilon(t, x, v) b(v). \end{aligned} \tag{3.7}$$

Since $\eta^\varepsilon(t, x, v, v')$ is antisymmetric with respect to the exchange of v and v' , by integrating (3.5) with respect to $\pi(dv)$ we deduce the continuity equation

$$\partial_t \rho^\varepsilon + \nabla \cdot j^\varepsilon = 0. \tag{3.8}$$

Theorem 3.3. Assume that $\rho_0^\varepsilon \rightarrow \rho_0$ in $\mathcal{P}(\mathbb{T}^d)$ and $\lim_{\varepsilon \rightarrow 0} \mathcal{H}(f_0^\varepsilon) = H(\rho_0)$. Then the sequence $(\rho^\varepsilon, j^\varepsilon)$ converges in $C([0, T]; \mathcal{P}(\mathbb{T}^d)) \times \mathcal{M}([0, T] \times \mathbb{T}^d; \mathbb{R}^d)$ to the solution to the heat equation as in Definition 2.1, with initial datum ρ_0 and diffusion coefficient

$$D = \pi(b \otimes (-\mathcal{L})^{-1}b). \tag{3.9}$$

Note that, by Assumption 3.1, b/λ and $\xi = (-\mathcal{L})^{-1}b$ are in $L^2(\tilde{\pi}; \mathbb{R}^d)$, hence the diffusion coefficient $D = \pi(\lambda) \tilde{\pi}((b/\lambda) \otimes \xi)$ is finite.

The proof of this theorem will be achieved according to the following strategy. We first show precompactness of the sequence $\{(\rho^\varepsilon, j^\varepsilon)\}$, we then consider a converging subsequence $(\rho^\varepsilon, j^\varepsilon) \rightarrow (\rho, j)$ and take the inferior limit in the inequality

(3.6). By the hypothesis of the theorem, $\mathcal{H}(f_0^\varepsilon) \rightarrow H(\rho_0)$ and we prove that the inferior limit of the left-hand side of (3.6) majorizes the left-hand side of (2.7). The statement follows by the uniqueness in Proposition 2.2. We introduce the following notations. If $(f^\varepsilon, \eta^\varepsilon)$ satisfy (3.5), (3.6), recalling (3.1), we set

$$\begin{aligned} f^\varepsilon(t, x, v) &= u_\varepsilon^2(t, x, v), \quad u_\varepsilon(t, x, v) = \bar{u}_\varepsilon(t, x) + \tilde{u}_\varepsilon(t, x, v), \\ \bar{u}_\varepsilon(t, x) &= \tilde{\pi}(u_\varepsilon(t, x, \cdot)). \end{aligned} \tag{3.10}$$

We will use the following bounds that hold uniformly for $t \in [0, T]$. By the Cauchy-Schwarz inequality

$$\int dx \bar{u}_\varepsilon^2(t) = \frac{1}{\pi(\lambda)^2} \int dx \left(\int d\pi \lambda \sqrt{f^\varepsilon(t)} \right)^2 \leq \frac{\pi(\lambda^2)}{\pi(\lambda)^2}. \tag{3.11}$$

Moreover, by the basic entropy inequality, for each $\gamma > 0$

$$\begin{aligned} \int dx \int d\pi f^\varepsilon(t) \frac{|b|^2}{\lambda} &\leq \frac{1}{\gamma} \mathcal{H}(f^\varepsilon(t)) + \frac{1}{\gamma} \log \pi \left(\exp \left\{ \gamma \frac{|b|^2}{\lambda} \right\} \right) \\ &\leq \frac{1}{\gamma} \mathcal{H}(f_0^\varepsilon) + \frac{1}{\gamma} \log \pi \left(\exp \left\{ \gamma \frac{|b|^2}{\lambda} \right\} \right). \end{aligned} \tag{3.12}$$

Lemma 3.4. *There exists a constant C such that for any $\varepsilon \in (0, 1)$*

$$\int dx \int d\pi \tilde{u}_\varepsilon(t)^2 \frac{|b|^2}{\lambda} < C, \quad t \in [0, T], \tag{3.13}$$

$$\frac{1}{\varepsilon^2} \int_0^T dt \int dx \int d\tilde{\pi} \tilde{u}_\varepsilon(t)^2 < C. \tag{3.14}$$

Proof. In order to prove (3.13), since $\tilde{u}_\varepsilon = \sqrt{f^\varepsilon} - \bar{u}_\varepsilon$, for each $t \in [0, T]$

$$\int dx \int d\tilde{\pi} \tilde{u}_\varepsilon(t)^2 \frac{|b|^2}{\lambda^2} \leq \frac{2}{\pi(\lambda)} \int dx \int d\pi f^\varepsilon(t) \frac{|b|^2}{\lambda} + \frac{2}{\pi(\lambda)} \int dx \bar{u}_\varepsilon^2(t) \int d\pi \frac{|b|^2}{\lambda}.$$

The first term on the right-hand side is bounded by (3.12), while the second term is bounded, since b^2/λ has finite exponential moments, by (3.11).

Regarding (3.14), by the Poincaré inequality (3.2),

$$\begin{aligned} &\frac{1}{\varepsilon^2} \int_0^T dt \int dx \int d\tilde{\pi} \tilde{u}_\varepsilon^2(t) \\ &\leq \frac{C_0}{\varepsilon^2} \int_0^T dt \int dx \iint \tilde{\pi}(dv) \tilde{\pi}(dv') \frac{\sigma(v, v')}{\lambda(v)\lambda(v')} [u_\varepsilon(t, x, v) - u_\varepsilon(t, x, v')]^2 \\ &\leq \frac{C_0}{\varepsilon^2} \pi(\lambda)^2 \int_0^T dt \mathcal{E}(f^\varepsilon(t)) \leq C \mathcal{H}(f_0^\varepsilon), \end{aligned}$$

which concludes the proof. □

Lemma 3.5. *The set $\{(\rho^\varepsilon, j^\varepsilon)\}_{\varepsilon \in (0,1]} \subset C([0, T]; \mathcal{P}(\mathbb{T}^d)) \times \mathcal{M}([0, T] \times \mathbb{T}^d; \mathbb{R}^d)$ is precompact.*

Proof. Given $0 \leq s < t \leq T$, the restriction of the measure $dJ^\varepsilon = j^\varepsilon dt dx$ to $[s, t] \times \mathbb{T}^d$ is denoted by $J_{s,t}^\varepsilon$. We will prove the following bound. There exists a constant C independent on s, t , such that for any $\varepsilon \in (0, 1]$ and any $w \in C([s, t] \times \mathbb{T}^d; \mathbb{R}^d)$

$$|J_{s,t}^\varepsilon(w)| = \left| \int_s^t d\tau \int dx j^\varepsilon(\tau, x) \cdot w(\tau, x) \right| \leq C \sqrt{t-s} \|w\|_\infty. \tag{3.15}$$

Let us first show that it implies the statement. Choosing $s = 0, t = T$, and applying the Banach-Alaoglu theorem, (3.15) directly yields the precompactness of $\{j^\varepsilon\}$. Since $\mathcal{P}(\mathbb{T}^d)$ is compact, by the Ascoli-Arzelà theorem, to prove precompactness of $\{\rho^\varepsilon\}$ it is enough to show that for each $\phi \in C^1(\mathbb{T}^d)$

$$\lim_{\delta \downarrow 0} \sup_{\varepsilon \in (0,1]} \sup_{\substack{t,s \in [0,T] \\ |t-s| < \delta}} \left| \int dx [\rho^\varepsilon(t, x) - \rho^\varepsilon(s, x)] \phi(x) \right| = 0. \tag{3.16}$$

From the continuity equation (3.8) we deduce

$$\int dx [\rho^\varepsilon(t, x) - \rho^\varepsilon(s, x)] \phi(x) = \int_s^t d\tau \int dx j^\varepsilon(\tau, x) \cdot \nabla \phi(x) = J_{s,t}^\varepsilon(\nabla \phi),$$

so that (3.16) follows readily from (3.15).

To prove (3.15), using the decomposition (3.10), since b has mean zero with respect to π , by definition (3.7) of j^ε we get

$$\begin{aligned} J_{s,t}^\varepsilon(w) &= \frac{1}{\varepsilon} \int_s^t d\tau \int dx \int \pi(dv) f^\varepsilon(\tau, x, v) b(v) \cdot w(\tau, x) \\ &= \frac{\pi(\lambda)}{\varepsilon} \int_s^t d\tau \int dx \int \tilde{\pi}(dv) [\tilde{u}_\varepsilon^2(\tau, x, v) + 2\bar{u}_\varepsilon(\tau, x) \tilde{u}_\varepsilon(\tau, x, v)] \frac{b(v)}{\lambda(v)} \\ &\quad \cdot w(\tau, x). \end{aligned}$$

By Young’s inequality, for each $\gamma > 0$

$$\begin{aligned} \frac{1}{\varepsilon} \tilde{u}_\varepsilon^2 \frac{|b|}{\lambda} &\leq \frac{\gamma}{2\varepsilon^2} \tilde{u}_\varepsilon^2 + \frac{1}{2\gamma} \tilde{u}_\varepsilon^2 \frac{|b|^2}{\lambda^2} \\ \frac{2}{\varepsilon} \bar{u}_\varepsilon |\tilde{u}_\varepsilon| \frac{|b|}{\lambda} &\leq \frac{\gamma}{\varepsilon^2} \tilde{u}_\varepsilon^2 + \frac{1}{\gamma} \tilde{u}_\varepsilon^2 \frac{|b|^2}{\lambda^2}. \end{aligned}$$

Then we obtain

$$\begin{aligned} |J_{s,t}^\varepsilon(w)| \leq \|w\|_\infty &\left\{ \frac{3\gamma}{2\varepsilon^2} \int_0^T d\tau \int dx \int d\tilde{\pi} \tilde{u}_\varepsilon^2 + \frac{1}{2\gamma} \int_s^t d\tau \int dx \int d\tilde{\pi} \tilde{u}_\varepsilon^2 \frac{|b|^2}{\lambda^2} \right. \\ &\left. + \frac{1}{\gamma} \int_s^t d\tau \int dx \tilde{u}_\varepsilon^2 \int d\tilde{\pi} \frac{|b|^2}{\lambda^2} \right\}. \end{aligned}$$

Using (3.14), (3.13), the fact that $|b|^2/\lambda$ has finite exponential moments, and (3.11), we obtain that there exists C such that

$$|J_{s,t}^\varepsilon(w)| \leq \frac{C}{2} \|w\|_\infty \left(\gamma + \frac{t-s}{\gamma} \right),$$

then (3.15) is obtained by choosing $\gamma = \sqrt{t-s}$. □

Lemma 3.6. *Assume that $\rho^\varepsilon \rightarrow \rho$. Then for each $t \in [0, T]$*

$$\varliminf_{\varepsilon \rightarrow 0} \mathcal{H}(f^\varepsilon(t)) \geq H(\rho(t)).$$

Proof. The statement is a direct consequence of the convexity and lower semicontinuity of the relative entropy. □

Lemma 3.7. *Assume that $(\rho^\varepsilon, j^\varepsilon) \rightarrow (\rho, j)$. Then*

$$\varliminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^2} \int_0^T dt \mathcal{E}(f^\varepsilon(t)) \geq R(\rho, j). \tag{3.17}$$

Proof. Assume first that condition (ii) in Assumption 3.2 holds. Recall (3.1) and observe that, in view of item (iii) in Assumption 3.1, $b/\lambda \in L^2(\tilde{\pi})$. Choose a sequence $\{a_n\}$, $a_n : \mathcal{V} \rightarrow \mathbb{R}^d$, converging to b/λ in $L^2(\tilde{\pi})$, such that: a_n is bounded, $\tilde{\pi}(a_n) = 0$ and $|a_n(v)| \leq |b(v)|/\lambda(v)$ for any $n \geq 1$. Upon extracting a subsequence, $a_n \rightarrow b/\lambda$ $\tilde{\pi}$ -a.e. Set $\omega_n := (\mathbb{I} - K)^{-1}a_n$. By the Poincaré inequality (3.2) $(\mathbb{I} - K)^{-1}$ is a bounded operator on the subspace of $L^2(\tilde{\pi})$ orthogonal to the constants; hence ω_n converges to $\xi = (\mathbb{I} - K)^{-1}(b/\lambda)$ in $L^2(\tilde{\pi})$. Moreover, by condition (ii) in Assumption 3.2, for each $n \geq 1$ ω_n is bounded.

Fix $w \in C([0, T] \times \mathbb{T}^d; \mathbb{R}^d)$. In the variational representation (2.19) for \mathcal{E} we chose the test function $\log(1 + \varepsilon w(t, x) \cdot \omega_n(v))$, with ε small enough, and we deduce

$$\frac{1}{2} \frac{1}{\varepsilon^2} \int_0^T dt \mathcal{E}(f^\varepsilon(t)) \geq \frac{1}{\varepsilon} \int_0^T dt \int dx \int d\pi f^\varepsilon \frac{w \cdot (-\mathcal{L})\omega_n}{1 + \varepsilon w \cdot \omega_n}.$$

Since ω_n is bounded, by Taylor expansion we obtain

$$\begin{aligned} \varliminf_{\varepsilon \rightarrow 0} \frac{1}{2} \frac{1}{\varepsilon^2} \int_0^T dt \mathcal{E}(f^\varepsilon(t)) &\geq \varliminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_0^T dt \int dx \int d\pi f^\varepsilon w \cdot (-\mathcal{L})\omega_n \\ &\quad - \varlimsup_{\varepsilon \rightarrow 0} \int_0^T dt \int dx \int d\pi f^\varepsilon w \cdot \omega_n w \cdot (-\mathcal{L})\omega_n. \end{aligned}$$

Regarding the first term on the right-hand side, we write

$$\frac{1}{\varepsilon} \int_0^T dt \int dx \int d\pi f^\varepsilon w \cdot (-\mathcal{L})\omega_n = \int_0^T dt \int dx j^\varepsilon \cdot w + A_{\varepsilon,n},$$

with

$$A_{\varepsilon,n} = \frac{1}{\varepsilon} \int_0^T dt \int dx \int d\pi f^\varepsilon w \cdot ((-\mathcal{L})\omega_n - b).$$

We will show that

$$\lim_{n \rightarrow \infty} \sup_{\varepsilon > 0} |A_{\varepsilon,n}| = 0 \tag{3.18}$$

and

$$\overline{\lim}_{n \rightarrow \infty} \overline{\lim}_{\varepsilon \rightarrow 0} \int_0^T dt \int dx \int d\pi f^\varepsilon w \cdot \omega_n w \cdot (-\mathcal{L})\omega_n \leq \int_0^T dt \int dx \rho w \cdot D w, \tag{3.19}$$

with D given by (3.9). Then, by optimizing over w and using the variational representation (2.5), the statement follows.

Postponing the proof of these two bounds, we consider the case that condition (i) in Assumption 3.2 holds. Fix $w \in C([0, T] \times \mathbb{T}^d; \mathbb{R}^d)$, then in the variational representation for \mathcal{E} we choose the test function $\log(1 + \varepsilon w(t, x) \cdot (-\mathcal{L})^{-1}b(v))$, with ε small enough. By Taylor expansion we deduce

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \frac{1}{2} \frac{1}{\varepsilon^2} \int_0^T dt \mathcal{E}(f^\varepsilon(t)) &\geq \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_0^T dt \int dx \int d\pi f^\varepsilon w \cdot b \\ &\quad - \overline{\lim}_{\varepsilon \rightarrow 0} \int_0^T dt \int dx \int d\pi f^\varepsilon w \cdot (-\mathcal{L})^{-1}b w \cdot b. \end{aligned}$$

Recalling (3.7) and the variational representation (2.5), it suffices to show

$$\overline{\lim}_{\varepsilon \rightarrow 0} \int_0^T dt \int dx \int d\pi f^\varepsilon w \cdot b w \cdot (-\mathcal{L})^{-1}b \leq \int_0^T dt \int dx \rho w \cdot D w. \tag{3.20} \quad \square$$

Proof of (3.18). According to the decomposition of f^ε in (3.10), we write

$$\begin{aligned} A_{\varepsilon,n} &= \frac{1}{\varepsilon} \int_0^T dt \int dx \int d\pi \tilde{u}_\varepsilon^2 w \cdot ((-\mathcal{L})\omega_n - b) \\ &\quad + \frac{2}{\varepsilon} \int_0^T dt \int dx \int d\pi \tilde{u}_\varepsilon \bar{u}_\varepsilon w \cdot ((-\mathcal{L})\omega_n - b), \end{aligned} \tag{3.21}$$

where we used that $\pi[(-\mathcal{L})\omega_n - b] = 0$. By Young’s inequality, for any $\gamma > 0$

$$\begin{aligned} \frac{1}{\varepsilon} \tilde{u}_\varepsilon^2 |w| \frac{1}{\lambda} |(-\mathcal{L})\omega_n - b| &\leq \frac{\gamma}{2\varepsilon^2} \tilde{u}_\varepsilon^2 + \frac{1}{2\gamma} |w|^2 \tilde{u}_\varepsilon^2 \frac{1}{\lambda^2} |(-\mathcal{L})\omega_n - b|^2, \\ \frac{2}{\varepsilon} \frac{1}{\lambda} \bar{u}_\varepsilon |\tilde{u}_\varepsilon| |w| |(-\mathcal{L})\omega_n - b| &\leq \frac{\gamma}{\varepsilon^2} \tilde{u}_\varepsilon^2 + \frac{1}{\gamma} |w|^2 \tilde{u}_\varepsilon^2 \frac{1}{\lambda^2} |(-\mathcal{L})\omega_n - b|^2. \end{aligned}$$

Then

$$\begin{aligned}
 |A_{\varepsilon,n}| &\leq \frac{3\gamma}{2\varepsilon^2} \int_0^T dt \int dx \int d\tilde{\pi} \tilde{u}_\varepsilon^2 \\
 &\quad + \frac{\pi[\lambda]^2}{2\gamma} \|w\|_\infty^2 \int_0^T dt \int dx \int d\tilde{\pi} \tilde{u}_\varepsilon^2 \frac{1}{\lambda^2} |((-\mathcal{L})\omega_n - b)|^2 \\
 &\quad + \frac{\pi[\lambda]^2}{\gamma} \|w\|_\infty^2 \int_0^T dt \int dx \tilde{u}_\varepsilon^2 \int d\tilde{\pi} \frac{1}{\lambda^2} |((-\mathcal{L})\omega_n - b)|^2.
 \end{aligned}
 \tag{3.22}$$

We claim that for each $\gamma > 0$ the second and the third term on the right-hand side vanishes as $n \rightarrow \infty$ uniformly in ε . Since the first term on the right-hand side can be bounded by using (3.14), we then conclude taking the limit $\gamma \rightarrow 0$.

To prove the claim, observe that, by construction of the sequence $\{\omega_n\}$,

$$\begin{aligned}
 [(-\mathcal{L})\omega_n](v) &= \lambda(v)a_n(v) \rightarrow b(v) \quad \pi\text{-a.e.} \\
 |(-\mathcal{L})\omega_n(v) - b(v)| &\leq 2|b(v)|.
 \end{aligned}
 \tag{3.23}$$

As $\int dx \tilde{u}_\varepsilon^2$ is bounded uniformly in ε by (3.11), we conclude by dominated convergence and (3.13). \square

Proof of (3.19). It is enough to show that for each n

$$\lim_{\varepsilon \rightarrow 0} \int_0^T dt \int dx \int d\pi (f^\varepsilon - \rho^\varepsilon) w \cdot \omega_n w \cdot (-\mathcal{L})\omega_n = 0.
 \tag{3.24}$$

Indeed, by construction of the sequence a_n

$$\lim_{n \rightarrow \infty} \pi(\omega_n \otimes (-\mathcal{L})\omega_n) = \lim_{n \rightarrow \infty} \pi(\lambda)\tilde{\pi}(\omega_n \otimes a_n) = \pi(\lambda)\tilde{\pi}(\xi \otimes \frac{b}{\lambda}) = D.$$

In order to prove (3.24), by using the decomposition of f^ε in (3.10),

$$f^\varepsilon - \rho^\varepsilon = 2\bar{u}_\varepsilon [\tilde{u}_\varepsilon - \pi(\tilde{u}_\varepsilon)] + \tilde{u}_\varepsilon^2 - \pi(\tilde{u}_\varepsilon^2).$$

Since ω_n is bounded and $|(-\mathcal{L})\omega_n(v)| \leq |b(v)|$, it suffices

$$\lim_{\varepsilon \rightarrow 0} \int_0^T dt \int dx \left\{ \bar{u}_\varepsilon \pi(|\tilde{u}_\varepsilon||b|) + \bar{u}_\varepsilon \pi(|\tilde{u}_\varepsilon|)\pi(|b|) + \pi(\tilde{u}_\varepsilon^2|b|) + \pi(\tilde{u}_\varepsilon^2)\pi(|b|) \right\} = 0.
 \tag{3.25}$$

By Cauchy-Schwarz, Lemma 3.4, and (3.11), we directly conclude that the first and third term vanishes as $\varepsilon \rightarrow 0$. To analyze the fourth term, given $\delta > 0$ we write

$$\int_0^T dt \int dx \pi(\tilde{u}_\varepsilon^2) = \int_0^T dt \int dx \pi(\tilde{u}_\varepsilon^2 \chi_{\{\lambda \geq \delta\}}) + \int_0^T dt \int dx \pi(\tilde{u}_\varepsilon^2 \chi_{\{\lambda < \delta\}}).$$

By (3.14), the first term on the right-hand side vanishes as $\varepsilon \rightarrow 0$. It is therefore enough to show that the second term vanishes as $\delta \rightarrow 0$ uniformly in ε . To this end, recalling that $\tilde{u}_\varepsilon = u_\varepsilon - \bar{u}_\varepsilon$ with $u_\varepsilon^2 = f^\varepsilon$,

$$\begin{aligned} \int dx \int d\pi \tilde{u}_\varepsilon^2 \chi_{\{\lambda < \delta\}} &= \int dx \int d\pi (u_\varepsilon - \bar{u}_\varepsilon)^2 \chi_{\{\lambda < \delta\}} \\ &\leq 2 \int dx \int d\pi f^\varepsilon \chi_{\{\lambda < \delta\}} + 2\pi(\lambda < \delta) \int dx \bar{u}_\varepsilon^2, \end{aligned}$$

and we conclude by using (3.11), the basic entropy inequality and the assumption $\pi(\lambda = 0) = 0$. To complete the proof of (3.25), we observe that by Schwartz inequality and (3.11) the previous argument also implies that the second term vanishes as $\varepsilon \rightarrow 0$. \square

Proof of (3.20). As before, it suffices to show that

$$\lim_{\varepsilon \rightarrow 0} \int_0^T dt \int dx \int d\pi [f^\varepsilon - \rho^\varepsilon] w \cdot b w \cdot (-\mathcal{L})^{-1} b = 0.$$

Since $(-\mathcal{L})^{-1} b$ is bounded, this follows from (3.25). \square

Lemma 3.8. *Assume that $(\rho^\varepsilon, j^\varepsilon) \rightarrow (\rho, j)$. Then*

$$\liminf_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^2} \mathcal{R}(f^\varepsilon, \eta^\varepsilon) \geq \int_0^T dt E(\rho(t)). \tag{3.26}$$

Proof. Assume first that condition (ii) in Assumption 3.2 holds. Let a_n and ω_n as in the previous lemma, and fix $\phi: (0, T) \times \mathbb{T}^d \rightarrow \mathbb{R}$ with compact support. In the variational formula (2.21) we choose $\alpha = 1$ and $\zeta(t, x, v, v') = \varepsilon \nabla \phi(t, x) \cdot (\omega_n(v') - \omega_n(v))$, then, by the antisymmetry of η^ε with respect to the exchange of v, v' ,

$$\begin{aligned} \frac{1}{\varepsilon^2} \mathcal{R}(f^\varepsilon, \eta^\varepsilon) &\geq -\frac{2}{\varepsilon} \int_0^T dt \int dx \nabla \phi(t, x) \cdot \iint \pi(dv) \pi(dv') \eta^\varepsilon(t, x, v, v') \omega_n(v) \\ &\quad - \frac{2}{\varepsilon^2} \int_0^T dt \int dx \iint \pi(dv) \pi(dv') f^\varepsilon(t, x, v) \sigma(v, v') \\ &\quad \times \{ \text{ch}(\varepsilon \nabla \phi(t, x) \cdot (\omega_n(v') - \omega_n(v))) - 1 \}. \end{aligned}$$

By the balance equation (3.5),

$$\begin{aligned} &-\frac{2}{\varepsilon} \int_0^T dt \int dx \nabla \phi(t, x) \cdot \iint \pi(dv) \pi(dv') \eta^\varepsilon(t, x, v, v') \omega_n(v) \\ &= -2\varepsilon \int_0^T dt \int dx \partial_t \nabla \phi \cdot \int d\pi f^\varepsilon \omega_n - 2 \int_0^T dt \int dx \int d\pi f^\varepsilon \nabla \cdot [\omega_n \cdot \nabla \phi b]. \end{aligned}$$

Since ω_n is bounded, the first term on the right-hand side above vanishes as $\varepsilon \rightarrow 0$. Therefore, by Taylor expansion of ch ,

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^2} \mathcal{R}(f^\varepsilon, \eta^\varepsilon) \\ & \geq - \overline{\lim}_{\varepsilon \rightarrow 0} 2 \int_0^T dt \int dx \int \pi(dv) f^\varepsilon(t, x, v) \nabla \cdot (\omega_n(v) \cdot \nabla \phi(t, x) b(v)) \\ & \quad - \overline{\lim}_{\varepsilon \rightarrow 0} \int_0^T dt \int dx \iint \pi(dv) \pi(dv') f^\varepsilon(t, x, v) \sigma(v, v') \\ & \quad \times \left[\nabla \phi(t, x) \cdot (\omega_n(v') - \omega_n(v)) \right]^2. \end{aligned}$$

We will show that

$$\overline{\lim}_{n \rightarrow \infty} \overline{\lim}_{\varepsilon \rightarrow 0} 2 \int_0^T dt \int dx \int d\pi f^\varepsilon \nabla \cdot [\omega_n \cdot \nabla \phi b] \leq 2 \int_0^T dt \int dx \rho \nabla \cdot [D \nabla \phi] \tag{3.27}$$

and

$$\begin{aligned} & \overline{\lim}_{n \rightarrow \infty} \overline{\lim}_{\varepsilon \rightarrow 0} \int_0^T dt \int dx \int \pi(dv) \pi(dv') f^\varepsilon(t, x, v) \sigma(v, v') \\ & \quad \times \left[\nabla \phi(t, x) \cdot (\omega_n(v') - \omega_n(v)) \right]^2 \tag{3.28} \\ & \leq 2 \int_0^T dt \int dx \rho(t, x) \nabla \phi(t, x) \cdot D \nabla \phi(t, x). \end{aligned}$$

Then, by optimizing over ϕ and using the variational representation (2.4), the statement follows.

Assume now that condition (i) in Assumption 3.2 holds. Then we choose as test functions $\alpha = 1$ and $\zeta(t, x, v, v') = \varepsilon \nabla \phi(t, x) \cdot (-\mathcal{L})^{-1}(b(v') - b(v))$, with a smooth $\phi: (0, T) \times \mathbb{T}^d \rightarrow \mathbb{R}^d$ with compact support. Using the fact that $(\mathcal{L})^{-1}b$ is bounded we repeat the same arguments as above and therefore we have to show that

$$\overline{\lim}_{\varepsilon \rightarrow 0} 2 \int_0^T dt \int dx \int d\pi f^\varepsilon \nabla \cdot [(-\mathcal{L})^{-1}b \cdot \nabla \phi b] \leq 2 \int_0^T dt \int dx \rho \nabla \cdot [D \nabla \phi] \tag{3.29}$$

and

$$\begin{aligned} & \overline{\lim}_{\varepsilon \rightarrow 0} \int_0^T dt \int dx \iint \pi(dv) \pi(dv') f^\varepsilon(t, x, v) \sigma(v, v') \\ & \quad \times \left[\nabla \phi(t, x) \cdot (-\mathcal{L})^{-1}(b(v') - b(v)) \right]^2 \tag{3.30} \\ & \leq 2 \int_0^T dt \int dx \rho(t, x) \nabla \phi(t, x) \cdot D \nabla \phi(t, x). \end{aligned}$$

□

Proof of (3.27). We claim that for each n

$$\lim_{\varepsilon \rightarrow 0} \int_0^T dt \int dx \int d\pi [f^\varepsilon - \rho^\varepsilon] \nabla \cdot [\omega_n \cdot \nabla \phi b] = 0,$$

which is proven exactly as (3.24) observing that there we used the bound $|(-\mathcal{L})\omega_n| \leq |b|$. Since $\rho^\varepsilon \rightarrow \rho$ and, by construction of the sequence $\{\omega_n\}$, $\lim_n \pi(\omega_n \otimes b) = D$, we then conclude. \square

Proof of (3.28). We first show that for each n

$$\begin{aligned} &\lim_{\varepsilon \rightarrow 0} \int_0^T dt \int dx \int \pi(dv) \pi(dv') [f^\varepsilon(t, x, v) - \rho^\varepsilon(t, x)] \sigma(v, v') \\ &\quad \times \left[\nabla \phi(t, x) \cdot (\omega_n(v') - \omega_n(v)) \right]^2 = 0. \end{aligned}$$

Since ω_n is bounded and $\lambda(v) = \int \pi(dv') \sigma(v, v')$, it is enough to prove that

$$\lim_{\varepsilon \rightarrow 0} \int_0^T dt \int dx \int d\pi \lambda |f^\varepsilon - \rho^\varepsilon| = 0,$$

whose proof is achieved by the same arguments used in the proof of (3.19). We then conclude by observing that

$$\begin{aligned} &\lim_{n \rightarrow \infty} \iint \pi(dv) \pi(dv') \sigma(v, v') (\omega_n(v') - \omega_n(v)) \otimes (\omega_n(v') - \omega_n(v)) \\ &= \lim_{n \rightarrow \infty} 2\pi(\omega_n \otimes (-\mathcal{L})\omega_n) = 2D. \end{aligned} \quad \square$$

Proofs of (3.29) and (3.30). These are achieved as the proofs of (3.27) and (3.28) with $(-\mathcal{L})^{-1}b$ instead of ω_n . \square

4. Specific examples

We consider three examples of linear Boltzmann equations and we show that they meet the requirements in the Assumptions 2.3, 3.1 and 3.2. For all of them the convergence to a diffusion is a classical result, therefore they are suitable testers for the machinery. We emphasize however that we do not require the initial distribution to be in L^2 , as it is usual in the classical approaches, but only with finite entropy with respect to the reference measure.

4.1. Boltzmann-Grad limit for the Lorentz gas with hard scatterers

The first example is the linear Boltzmann equation derived for the one particle distribution in [20], starting from the Lorentz gas moving in a random array of fixed scatterers (hard spheres), in the Boltzmann-Grad limit. Since collisions are elastic, the kinetic energy is preserved, therefore the phase space is $\mathbb{T}^d \times S_{|v|}^{d-1}$, $d \geq 2$, where $S_{|v|}^{d-1}$ is the d -dimensional sphere with radius $|v|$. Without loss of generality, we assume $|v| = 1$. The equation then reads

$$\partial_t f(t, x, v) + v \cdot \nabla_x f(t, x, v) = \int_{S^{d-1}} d\hat{n} [\hat{n} \cdot v]_+ [f(t, x, v') - f(t, x, v)], \tag{4.1}$$

where $v' = v - 2(v \cdot \hat{n})\hat{n}$. The invariant measure is the uniform measure $d\hat{n}$ on S^{d-1} , the scattering rate λ is equal to c , for some constant c depending on the dimension d . In order to identify the scattering kernel, we consider here the case $d = 2$, referring to Appendix B for analogous computations if $d \geq 3$. By identifying the velocity $v \in S^1$ with the angle θ , we rewrite the previous equation as

$$\partial_t f(t, x, \theta) + b(\theta) \cdot \nabla_x f(t, x, \theta) = \frac{1}{2} \int_{S^1} d\theta' \left| \sin \frac{\theta - \theta'}{2} \right| [f(t, x, \theta') - f(t, x, \theta)],$$

with $b(\theta) = (\cos \theta, \sin \theta)$. In particular, the scattering kernel is $\sigma(\theta, \theta') = \left| \sin \frac{\theta - \theta'}{2} \right|$. Recalling the definition (3.1) of $\tilde{\pi}$, the operator K with kernel σ/λ is compact in $L^2(\tilde{\pi})$. Since 1 is a simple eigenvalue of K , then the modified chain has spectral gap. Moreover, as shown in [6, Lemma 4.1], $L^{-1}v$ is bounded. Hence Assumptions 2.3, 3.1 and the alternative (i) of Assumption 3.2 hold.

4.2. Rayleigh-Boltzmann equation

The Rayleigh-Boltzmann equation, also known as linear Boltzmann equation or Lorentz-Boltzmann equation, has been derived in the Boltzmann-Grad limit by looking at the distribution of a tracer particle in a gas of particles (hard spheres) in thermal equilibrium [8]. The velocity space is then \mathbb{R}^d , $d \geq 2$, and the reference measure is the Maxwellian distribution with temperature β^{-1} , whose density with respect to the Lebesgue measure is denoted by h_β . The equation reads

$$\begin{aligned} & \partial_t f(t, x, v) + v \cdot \nabla_x f(t, x, v) \\ &= \int_{\mathbb{R}^d} dv_1 h_\beta(v_1) \int_{S^{d-1}} d\hat{n} [\hat{n} \cdot (v - v_1)]_+ [f(t, x, v') - f(t, x, v)], \end{aligned} \tag{4.2}$$

where $v' = v - \hat{n} \cdot (v - v_1)\hat{n}$. As shown in the Appendix B, the scattering rate is

$$\lambda(v) = \chi \int_{\mathbb{R}^d} dv_1 h_\beta(v_1) |v - v_1|,$$

where χ is the constant given by $\chi = \int_{S^{d-1}} d\hat{n} [\hat{n} \cdot \hat{v}]_+$, in which $\hat{v} \in S^{d-1}$. In particular λ is bounded away from 0 and it has linear growth for large $|v|$. Therefore $|v|$ and $|v|^2/\lambda(v)$ have all the exponential moments with respect to $h_\beta(v) dv$. In Appendix B we identify the scattering kernel σ , see (B.2). From this expression and the properties of λ , recalling the definition (3.1), it follows that K in (3.3) has a kernel in $L^2(\tilde{\pi} \times \tilde{\pi})$. Hence K is compact in $L^2(\tilde{\pi})$. Since 1 is simple eigenvalue of K , then $(\mathbf{I} - K)$ has spectral gap. The previous statements imply that Assumptions 2.3 and 3.1 hold. The proof of alternative (i) in Assumption 3.2 is the content of Appendix B.

4.3. Linear phonon Boltzmann equation

The equation has been derived in the kinetic limit starting from an harmonic chain of oscillator perturbed by a stochastic conservative noise [7]. It describes the evolution of the energy density of the normal modes, or phonons, identified by a wave-number $k \in \mathbb{T}^d$. The velocity space is then \mathbb{T}^d . Let ω be the dispersion relation of the harmonic lattice, *i.e.* $\omega(k) = (v + 4 \sum_{i=1}^d \sin^2(\pi k_i))^{1/2}$, where $v \geq 0$ is the intensity of the pinning. A phonon with wave-number k travels with velocity $\frac{1}{2\pi} \nabla \omega$, then it is scattered. The corresponding Fokker-Planck equation reads

$$\partial_t f(t, x, k) + \frac{1}{2\pi} \nabla \omega(k) \cdot \nabla_x f(t, x, k) = \int_{\mathbb{T}^d} dk' \sigma(k, k') [f(t, x, k') - f(t, x, k)], \quad (4.3)$$

where the scattering kernel has the form $\sigma(k, k') = \sum_{i=1}^d \sin^2(\pi k_i) \sin^2(\pi k'_i)$ for $d \geq 2$. In dimension one it has a slightly different form, but despite the details the main features are that σ is positive, bounded and symmetric in the exchange k, k' . Then the invariant measure π is the Haar measure on the torus. Moreover σ vanishes in zero, since $\sigma(k, k') \sim |k|^2$ for small k , and the scattering rate λ has the same behavior. More precisely, $\lambda = c \sum_{i=1}^d \sin^2(\pi k_i)$, for some constant $c > 0$. Since $\partial_i \omega(k) = 2\pi \sin(2\pi k_i) / \omega(k)$, $i = 1, \dots, d$, in order to guarantee that $|\nabla \omega|^2 / \lambda$ has exponential moments we have to restrict to the case $v > 0$. This corresponds to assume that in the underlying harmonic chain the translational symmetry is broken, due to the presence of a on-site potential (pinning). Recalling the definition (3.1) of $\tilde{\pi}$, by the properties of σ and λ we deduce that the operator K defined in (3.3) has a kernel of the form $p(k, k') \tilde{\pi}(dk')$, with p strictly positive and bounded. Then K is a compact operator in $L^2(\tilde{\pi})$, and since 1 is a simple eigenvalue, then the modified chain has spectral gap. The previous statements imply that Assumptions 2.3 and 3.1 hold. Finally, since the modified chain satisfies the Doeblin condition, then for each $f \in L^\infty(\mathbb{T}^d)$ such that $\tilde{\pi}(f) = 0$ we have $\| \sum_{n \geq 0} K^n f \|_\infty \leq c \| f \|_\infty$, which implies alternative (ii) in Assumption 3.2.

In the unpinning case $v = 0$ the diffusion coefficient D diverges in dimension $d = 1, 2$. In these cases the asymptotics of the linear phonon Boltzmann equation is in fact a superdiffusion when $d = 1$ [5, 22] and a diffusion under an anomalous scaling, *i.e.* with logarithmic corrections, when $d = 2$ [4]; see also [29] for other

models with super-diffusive behavior. On the other hand, for $d \geq 3$ the diffusion coefficient D is finite even if $|b|^2/\lambda$ does not have exponential moments. Moreover, as discussed in [4], if the initial distribution satisfies suitable integrability conditions, the diffusive scaling holds. As the gradient flow approach here introduced requires only an entropy bound on the initial condition, it does not cover this case. It is not clear if this is just a limitation of the present approach or the diffusive limit fails if the integrability conditions are not satisfied. Indeed, phonons with small wave number are responsible of the ballistic transport which, in dimension one and two, induces the superdiffusion. If the initial conditions gives enough weight to those phonons, similar effects might occur also for $d \geq 3$.

Appendix

A. Entropy balance

We here prove (2.24). We can assume that its left-hand side is finite. Using also that P_r has bounded entropy for each $r \in [s, t]$ we deduce that $P_r(dx, dv) = f_r(x, v) dx \pi(dv)$ and $\Theta_{[s,t]}(dr, dx, dv, dv') = \eta_r(x, v, v') dr dx \pi(dv)\pi(dv')$. Moreover, recalling the function Ψ_κ in (2.14),

$$\begin{aligned} & \int_s^t dr \mathcal{E}(P_r) \\ &= \int_s^t dr \int dx \iint \pi(dv)\pi(dv') \sigma(v, v') \left[\sqrt{f_r(x, v')} - \sqrt{f_r(x, v)} \right]^2 \\ & \mathcal{R}^{s,t}(P, \Theta_{[s,t]}) \\ &= \int_s^t dr \int dx \iint \pi(dv)\pi(dv') \Psi_{\sigma(v,v')}(f_r(x, v), f_r(x, v'); \eta_r(x, v, v')) \end{aligned}$$

and

$$\int_s^t dr \mathcal{E}(P_r) + \mathcal{R}^{s,t}(P, \Theta_{[s,t]}) < +\infty. \tag{A.1}$$

We claim that, for f and η as above, the following entropy balance holds

$$\begin{aligned} & \mathcal{H}(P_t) - \mathcal{H}(P_{s'}) \\ &= \frac{1}{2} \int_{s'}^t dr \int dx \iint \pi(dv)\pi(dv') \eta_r(x, v, v') \left[\log f_r(x, v') - \log f_r(x, v) \right] \end{aligned} \tag{A.2}$$

for any $0 \leq s < s' < t \leq T$. Note that the last term is well defined by Legendre duality and (A.1). Informally, it is deduced by choosing the test function $\mathbb{1}_{[s',t]}(r) \log f_r(x, v)$ in the balance equation (2.22). The actual proof is carried out by a truncation argument that is next detailed.

Step 1. Approximation by space time convolutions. For $n \in \mathbb{N}$ let now $\chi_n: \mathbb{R} \rightarrow \mathbb{R}_+$ be a smooth approximation of the identity with compact support contained in the positive axis, and $g_n: \mathbb{T}^d \rightarrow \mathbb{R}_+$ be a smooth approximation of the identity. For $0 \leq s < s' \leq r \leq t \leq T$, by choosing n such that the $\text{supp}\chi_n \subset [0, s' - s]$, we define

$$f_r^n(x, v) := \int dr' \int dy \chi_n(r - r') g_n(x - y) f_{r'}(y, v)$$

$$\eta_r^n(x, v, v') := \int dr' \int dy \chi_n(r - r') g_n(x - y) \eta_{r'}(y, v, v').$$

As simple to check, the pair (f^n, η^n) satisfies the balance equation and, by (A.1) and convexity, there exists a constant C such that

$$\int_{s'}^t dr \int dx \iint \pi(dv)\pi(dv') \sigma(v, v') \left[\sqrt{f_r^n(x, v')} - \sqrt{f_r^n(x, v)} \right]^2$$

$$+ \int_{s'}^t dr \int dx \iint \pi(dv)\pi(dv') \Psi_{\sigma(v, v')}(f_r^n(x, v), f_r^n(x, v'); \eta_r^n(x, v, v')) \leq C,$$

and

$$\sup_{r \in [s', t]} \mathcal{H}(f_r^n) \leq \sup_{r \in [0, T]} \mathcal{H}(f_r) \leq C.$$

Step 2. Truncation of log. The balance equation (2.22) implies

$$\int dx \int \pi(dv) f_t^n(x, v) \phi(t, x, v) - \int dx \int \pi(dv) f_{s'}^n(x, v) \phi(s', x, v)$$

$$- \int_{s'}^t dr \int dx \int \pi(dv) f_r^n(x, v) \{ \partial_r \phi(r, x, v) + b(v) \cdot \nabla \phi(r, x, v) \} \quad (\text{A.3})$$

$$= \frac{1}{2} \int_{s'}^t dr \int dx \iint \pi(dv)\pi(dv') \eta^n(r, x, v, v') [\phi(r, x, v') - \phi(r, x, v)]$$

for all continuous functions $\phi: [s', t] \times \mathbb{T}^d \times \mathcal{V}$ with compact support in \mathcal{V} and continuously differentiable in the first two variables. Recalling that b has exponential moments, since f^n has bounded entropy and $\eta^n \in L^1$ we can use ϕ bounded instead of compactly supported.

Given $0 < \delta < L$ set

$$\log_{\delta, L}(u) = \begin{cases} \log \delta & \text{if } 0 < u < \delta \\ \log u & \text{if } \delta \leq u \leq L \\ \log L & \text{if } u > L. \end{cases} \quad (\text{A.4})$$

By a straightforward approximation we can choose as test function in (A.3) $\phi = \log_{\delta, L}(f^n)$, obtaining

$$\begin{aligned} & \int dx \int \pi(dv) f_t^n(x, v) \log_{\delta, L}(f_t^n(x, v)) \\ & - \int dx \int \pi(dv) f_{s'}^n(x, v) \log_{\delta, L}(f_{s'}^n(x, v)) \\ & - \int_{s'}^t dr \int dx \int \pi(dv) \mathbb{I}_{[\delta, L]}(f_r^n(x, v)) \{ \partial_r f_r^n(x, v) + b(v) \cdot \nabla f_r^n(x, v) \} \quad (\text{A.5}) \\ & = \frac{1}{2} \int_{s'}^t dr \int dx \iint \pi(dv) \pi(dv') \eta^n(r, x, v, v') \\ & \quad \times [\log_{\delta, L}(f_r^n(x, v)) - \log_{\delta, L}(f_r^n(x, v'))]. \end{aligned}$$

We observe that

$$\begin{aligned} & \int_{s'}^t dr \int dx \int \pi(dv) \mathbb{I}_{[\delta, L]}(f_r^n(x, v)) \{ \partial_r f_r^n(x, v) + b(v) \cdot \nabla f_r^n(x, v) \} \quad (\text{A.6}) \\ & = \int dx \int \pi(dv) (f_t^n(x, v) \wedge \delta) \vee L - \int dx \int \pi(dv) (f_{s'}^n(x, v) \wedge \delta) \vee L. \end{aligned}$$

Step 3. Removing the convolution. Since $\log_{\delta, L}$ is bounded, by dominated convergence we can remove regularization in space and time and we obtain

$$\begin{aligned} & \int dx \int \pi(dv) f_t(x, v) \log_{\delta, L}(f_t(x, v)) \\ & - \int dx \int \pi(dv) f_{s'}(x, v) \log_{\delta, L}(f_{s'}(x, v)) \\ & - \int dx \int \pi(dv) (f_t(x, v) \wedge \delta) \vee L \quad (\text{A.7}) \\ & + \int dx \int \pi(dv) (f_{s'}(x, v) \wedge \delta) \vee L. \\ & = \frac{1}{2} \int_{s'}^t dr \int dx \iint \pi(dv) \pi(dv') \eta(r, x, v, v') \\ & \quad \times [\log_{\delta, L}(f_r(x, v)) - \log_{\delta, L}(f_r(x, v'))]. \end{aligned}$$

Step 3. Removing the truncation of log. Here we take the limit $\delta \downarrow 0$ and $L \uparrow +\infty$ in (A.7). For the left-hand side this is accomplished by monotone convergence, in particular it converges to $\mathcal{H}(P_t) - \mathcal{H}(P_{s'})$. For the right-hand side, it is enough to show that

$$\begin{aligned} & \frac{1}{2} \int_{s'}^t dr \int dx \iint \pi(dv) \pi(dv') \eta(r, x, v, v') \quad (\text{A.8}) \\ & \times \left\{ [\log_{\delta, L}(f_r(x, v)) - \log(f_r(x, v))] - [\log_{\delta, L}(f_r(x, v')) - \log(f_r(x, v'))] \right\} \end{aligned}$$

vanishes as $\delta \downarrow 0$ and $L \uparrow +\infty$. We apply Young inequality in the form

$$pq \leq \psi_\alpha(p) + \psi_\alpha^*(q),$$

where, for $\alpha \geq 0$,

$$\psi_\alpha(p) = p \operatorname{ash} \frac{p}{\alpha} - \sqrt{p^2 + \alpha^2} + \alpha, \quad \psi_\alpha^*(q) = \alpha(\cosh q - 1).$$

Observe that ψ_α and ψ_α^* are even. By choosing $\alpha = 2\sigma(v, v')\sqrt{f_r(x, v)f_r(x, v')}$, $p = \eta_r(x, v, v')$ and $q = \frac{1}{2}[\log_{\delta, L}(f_r(x, v)) - \log(f_r(x, v)) - \log_{\delta, L}(f_r(x, v')) + \log(f_r(x, v'))]$, the first term is

$$\begin{aligned} & \int_{s'}^t dr \int dx \iint \pi(dv)\pi(dv')\psi_\alpha(\eta)[1 - \mathbb{I}_{[\delta, L]}(f_r(x, v))] \\ & \leq \int_{s'}^t dr \int dx \iint \pi(dv)\pi(dv') \Psi_{\sigma(v, v')}(f_r(x, v), f_r(x, v'); \eta_r(x, v, v')), \end{aligned}$$

which vanishes by dominated convergence since $\mathcal{R}(f, \eta)$ is finite. The second term has the following expression:

$$\begin{aligned} & \int_{s'}^t dr \int dx \iint \pi(dv)\pi(dv')\psi_\alpha^*(q) \\ = & \int_{s'}^t dr \int dx \iint \pi(dv)\pi(dv')[\mathbb{I}_{[0, \delta]}(f_r(x, v))\mathbb{I}_{[0, \delta]}(f_r(x, v')) \\ & + \mathbb{I}_{(L, +\infty)}(f_r(x, v))\mathbb{I}_{(L, +\infty)}(f_r(x, v'))]\sigma(v, v')(\sqrt{f_r(x, v)} - \sqrt{f_r(x, v')})^2 \\ & + 2 \int_{s'}^t dr \int dx \iint \pi(dv)\pi(dv')\mathbb{I}_{[0, \delta]}(f_r(x, v))\mathbb{I}_{[\delta, L]}(f_r(x, v')) \\ & \times \sigma(v, v')\sqrt{f_r(x, v)f_r(x, v')}\left(\frac{\sqrt{\delta}}{\sqrt{f_r(x, v)}} + \frac{\sqrt{f_r(x, v)}}{\sqrt{\delta}} - 2\right) \\ & + 2 \int_{s'}^t dr \int dx \iint \pi(dv)\pi(dv')\mathbb{I}_{[0, \delta]}(f_r(x, v))\mathbb{I}_{(L, +\infty)}(f_r(x, v')) \\ & \times \sigma(v, v')\sqrt{f_r(x, v)f_r(x, v')}\left(\frac{\sqrt{Lf_r(x, v)}}{\sqrt{\delta f_r(x, v')}} + \frac{\sqrt{\delta f_r(x, v')}}{\sqrt{Lf_r(x, v)}} - 2\right) \\ & + 2 \int_{s'}^t dr \int dx \iint \pi(dv)\pi(dv')\mathbb{I}_{[\delta, L]}(f_r(x, v))\mathbb{I}_{(L, +\infty)}(f_r(x, v')) \\ & \times \sigma(v, v')\sqrt{f_r(x, v)f_r(x, v')}\left(\frac{\sqrt{L}}{\sqrt{f_r(x, v)}} + \frac{\sqrt{f_r(x, v')}}{\sqrt{L}} - 2\right). \end{aligned}$$

We observe that the first integral on the right-hand side vanishes as $\delta \downarrow 0, L \uparrow +\infty$ since $\int_{s'}^t dr \mathcal{E}(f_r) < +\infty$. The term in the second integral is bounded by

$\sigma(v, v')\sqrt{\delta}\sqrt{f_r(x, v')}$ and the term in the third is bounded by $\sigma(v, v')\sqrt{\delta/L}f_r(x, v')$, then the two integrals vanish as $\delta \downarrow 0, L \uparrow +\infty$ since the scattering rate λ has all exponential moments and f_r has finite entropy. Finally, the term in the last integral is bounded by $\sigma(v, v')f_r(x, v')\mathbf{1}_{(L, +\infty)}(f_r(x, v'))$, which vanishes by dominated convergence.

Now we show that (2.24) holds. For $s' > s \geq 0$ it follows by applying again Young inequality with $p = -\eta_r(x, v, v')$ and $q = \frac{1}{2}[\log(f_r(x, v)) - \log(f_r(x, v'))]$ with the entropy balance (A.2). Finally the case $s' = s$ is achieved by the lower semi-continuity of \mathcal{H} .

B. Bounds on $(-\mathcal{L})^{-1}b$ for the Rayleigh gas

We identify the scattering kernel σ on the right-hand side of (4.2). Setting $z = v - v_1$, the collision operator in (4.2) becomes

$$\mathcal{L}f(v) = \int_{\mathbb{R}^d} dz h_\beta(v - z) \int_{S^{d-1}} d\hat{n} [\hat{n} \cdot z]_+ \{f(v') - f(v)\}, \tag{B.1}$$

where

$$v' = v - (\hat{n} \cdot z) \hat{n}.$$

Fixed \hat{n} , we can write $z = \alpha \hat{n} + z^\perp$, where z^\perp lies in the hyperplane of dimension $d - 1$ orthogonal to \hat{n} . We indicate with $v^\perp = v - (\hat{n} \cdot v) \hat{n}$, the projection of v on this hyperplane. We have $[\hat{n} \cdot z]_+ = [\alpha]_+, dz = d\alpha dz^\perp$, and

$$\mathcal{L}f(v) = \int_{S^{d-1}} d\hat{n} \int dz^\perp h_\beta^{d-1}(v^\perp - z^\perp) \int_0^{+\infty} d\alpha \alpha h_\beta^1(v \cdot \hat{n} - \alpha) \{f(v') - f(v)\},$$

where h_β^k is the Maxwellian distribution in dimension k , and now $v' = v - \alpha \hat{n}$. The integral in dz^\perp gives 1. Choosing the new variable $w = v - \alpha \hat{n}$, we have $\alpha = |v - w|, v \cdot \hat{n} = v \cdot (v - w)/|v - w|$,

$$v \cdot \hat{n} - \alpha = v \cdot (v - w)/|v - w| - |v - w| = w \cdot (v - w)/|v - w|$$

and $dw = \alpha^{d-1} d\alpha d\hat{n} = |v - w|^{d-1} d\alpha d\hat{n}$. Then

$$\mathcal{L}f(v) = \int_{\mathbb{R}^d} dw h_\beta^1(w \cdot (v - w)/|v - w|) \frac{1}{|v - w|^{d-2}} \{f(w) - f(v)\},$$

which is of the form (2.20) with

$$\begin{aligned} \sigma(v, w) &= \frac{1}{|v - w|^{d-2}} \frac{h_\beta^1(w \cdot (v - w)/|v - w|)}{h_\beta(w)} \\ &= \left(\frac{\beta}{2\pi}\right)^{\frac{1-d}{2}} \frac{1}{|v - w|^{d-2}} \exp\left\{\frac{\beta}{2} \frac{|w|^2|v - w|^2 - (w \cdot (v - w))^2}{|v - w|^2}\right\} \\ &= \left(\frac{\beta}{2\pi}\right)^{\frac{1-d}{2}} \frac{1}{|v - w|^{d-2}} \exp\left\{\frac{\beta}{2} \frac{|w|^2|v|^2 - (w \cdot v)^2}{|v - w|^2}\right\}. \end{aligned} \tag{B.2}$$

which is symmetric. We remark that for $d = 3$ this expression has been obtained in [26].

In order to prove that $\xi = -\mathcal{L}^{-1}v$ is bounded, we decompose \mathcal{L} in the gain and loss terms

$$-(\mathcal{L}f)(v) = \lambda(v)f(v) - (Gf)(v),$$

where

$$\lambda(v) = (G1)(v) = \int_{\mathbb{R}^d} dw h_\beta(w)\sigma(v, w) = \chi \int_{\mathbb{R}^d} dv_1 h_\beta(v_1)|v - v_1| = \lambda(|v|),$$

and $\chi = \int_{S^{d-1}} d\hat{n} [\hat{n} \cdot \hat{v}]_+$ for any unit vector \hat{v} . Observe that $(Gf)(v) = \lambda(v)(Kf)(v)$, with K defined in (3.3). Note that, for convexity,

$$\lambda(v) \geq \chi|v|. \tag{B.3}$$

We search for a bounded function $\gamma(|v|)$ such that $\xi(v) = \hat{v}\gamma(|v|)$. Then we have

$$\begin{aligned} (G\xi)(v) &= \int_{\mathbb{R}^d} dw h_\beta(w)\sigma(v, w)\gamma(|w|)\frac{w}{|w|} \\ &= \int_{\mathbb{R}^d} dw h_\beta(w)\sigma(v, w)\frac{\gamma(|w|)}{|w|} [(w - (\hat{v} \cdot w)\hat{v}) + (\hat{v} \cdot w)\hat{v}] \end{aligned}$$

where in the last step we decomposed w into the component along \hat{v} and the orthogonal part $w^\perp = w - (\hat{v} \cdot w)\hat{v}$. Since $|w|$ and $\sigma(v, w)$ are invariant in the exchange $w^\perp \rightarrow -w^\perp$, then

$$(G\xi)(v) = \hat{v} \int_{\mathbb{R}^d} dw h_\beta(w)\sigma(v, w)(\hat{v} \cdot \hat{w})\gamma(|w|).$$

Since the integral is invariant under rotations of v , we can define the operator \tilde{G} acting on functions on the positive half line by

$$(\tilde{G}f)(\rho) := \int_{\mathbb{R}^d} dw h_\beta(w)\sigma(\rho\hat{v}, w)(\hat{v} \cdot \hat{w})f(|w|),$$

so that, for $\xi(v) = \hat{v}\gamma(|v|)$, $(G\xi)(v) = \hat{v}(\tilde{G}\gamma)(|v|)$. As

$$(\tilde{G}f)(\rho) = \int_{w \cdot \hat{v} > 0} dw h_\beta(w)(\sigma(\rho\hat{v}, w) - \sigma(\rho\hat{v}, -w))(\hat{v} \cdot w)\gamma(|w|)$$

and, if $w \cdot v > 0$, then $\sigma(v, w) \geq \sigma(v, -w)$, the operator \tilde{G} has positive kernel.

Setting $\eta(\rho) = \lambda(\rho)\gamma(\rho)$, we look for the solution of the equation

$$\rho = \eta(\rho) - (A\eta)(\rho), \quad \rho \in \mathbb{R}^+, \tag{B.4}$$

where

$$\begin{aligned}
 (A\eta)(\rho) &= \left(\tilde{G} \frac{\eta}{\lambda} \right) (\rho) = \int_{\mathbb{R}^d} dw h_\beta(w) \sigma(\rho \hat{v}, w) (\hat{v} \cdot \hat{w}) \frac{\eta(|w|)}{\lambda(|w|)} \\
 &= \int_{\mathbb{R}^d} dv_1 h_\beta(v_1) \int_{S^{d-1}} d\hat{n} [\hat{n} \cdot (\rho \hat{v} - v_1)]_+ (\hat{v} \cdot \hat{v}') \frac{\eta(|v'|)}{\lambda(|v'|)},
 \end{aligned}
 \tag{B.5}$$

in which $v' = \rho \hat{v} - (\hat{n} \cdot (\rho \hat{v} - v_1)) \hat{n}$. The operator A is self-adjoint with respect to the scalar product

$$(f, g) = \int_{\mathbb{R}^d} dv h_\beta(v) \frac{1}{\lambda(|v|)} f(|v|) g(|v|),$$

defined for $f, g : [0, \infty) \rightarrow \mathbb{R}$. From the positivity of the kernel of the operator, it follows that if $f(\rho) \geq g(\rho)$ for any $\rho \geq 0$, then $(Af)(\rho) \geq (Ag)(\rho)$ for any $\rho \geq 0$. Moreover, if η is continuous in $[0, \infty)$, $(A\eta)$ is continuous in $[0, \infty)$ as follow from (B.5).

By definition of $\lambda(\rho)$

$$(A\lambda)(\rho) = \int_{\mathbb{R}^d} dw h_\beta(w) \sigma(\rho \hat{v}, w) (\hat{v} \cdot \hat{w}) < \lambda(\rho)$$

and the inequality is strict for any ρ because $\hat{v} \cdot \hat{w} < 1$ in a set of full measure. Observe that, using the definition of v' ,

$$\hat{v} \cdot \hat{v}' = \frac{\rho(1 - (\hat{n} \cdot \hat{v})^2) + (\hat{n} \cdot v_1)(\hat{n} \cdot \hat{v})}{\sqrt{\rho^2(1 - (\hat{n} \cdot \hat{v})^2) + (\hat{n} \cdot v_1)^2}},$$

which for fixed v_1 converges to $\sqrt{1 - (\hat{n} \cdot \hat{v})^2}$ when $\rho \rightarrow +\infty$, while $[\hat{n} \cdot (\rho \hat{v} - v_1)]_+ / \rho \rightarrow |\hat{n} \cdot \hat{v}|_+$. By dominated convergence

$$\lim_{\rho \rightarrow +\infty} \frac{1}{\rho} (A\lambda)(\rho) = \int_{\mathbb{R}^d} dv_1 h_\beta(v_1) \int_{S^{d-1}} d\hat{n} [\hat{n} \cdot \hat{v}]_+ \sqrt{1 - (\hat{n} \cdot \hat{v})^2} < \chi.$$

Since $\lim_{\rho \rightarrow +\infty} \lambda(\rho) / \rho = \int_{S^{d-1}} d\hat{n} [\hat{n} \cdot \hat{v}]_+ = \chi$, we then conclude that there exists a constant $0 < z < 1$ such that

$$(A\lambda)(\rho) < z\lambda(\rho)$$

for any $\rho \geq 0$. Since $\lambda(\rho) \geq \chi\rho$ (see (B.3)), if $\zeta = \chi(1 - z)$, then

$$\lambda(\rho) \geq \zeta\rho + (A\lambda)(\rho).$$

Denoting by $\text{id} : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ the identity function $\text{id}(\rho) = \rho$, and iterating the above expression, we get

$$\lambda \geq \zeta \text{id} + \zeta A \text{id} + A^2 \lambda \geq \dots \geq \zeta \sum_{k=0}^n A^k \text{id} + A^{n+1} \lambda,$$

which implies that $\eta = \sum_{k=0}^{+\infty} A^k \text{id}$ is a well defined, positive, function, and bounded by λ / ζ . Since η solves (B.4), then $\xi(v) = \hat{v} \eta(|v|) / \lambda(|v|)$, which is bounded by $1 / \zeta$.

References

- [1] ADAMS, N. DIRR, M. A. PELETIER and J. ZIMMER, *From a large-deviations principle to the Wasserstein gradient flow: a new micro-macro passage*, *Comm. Math. Phys.* **307** (2011), 791–815.
- [2] L. AMBROSIO, N. GIGLI and G. SAVARÉ, “Gradient Flows in Metric Spaces and in the Spaces of Probability Measures”, *Lectures in Mathematics*, ETH Zurich, Birkhäuser, 2005.
- [3] C. BARDOS, R. SANTOS and R. SENTIS, *Diffusion approximation and computation of the critical size*, *Trans. Amer. Math. Soc.* **284** (1984), 617–649.
- [4] G. BASILE, *From a kinetic equation to a diffusion under an anomalous scaling*, *Ann. Inst. H. Poincaré Probab. Stat.* **50** (2014), 1301–1322.
- [5] G. BASILE and A. BOVIER, *Convergence of a kinetic equation to a fractional diffusion equation*, *Markov Process. Related Fields* **16** (2010), 15–44.
- [6] G. BASILE, A. NOTA, F. PEZZOTTI and M. PULVIRENTI, *Derivation of the Fick’s Law for the Lorentz model in a low density regime*, *Comm. Math. Phys.* **336** (2015), 1607–1636.
- [7] G. BASILE, S. OLLA and H. SPOHN, *Energy transport in stochastically perturbed lattice dynamics*, *Arch. Ration. Mech. Anal.* **195** (2010), 171–203.
- [8] H. VAN BEIJEREN, O. E. LANFORD III, J. L. LEBOWITZ and H. SPOHN, *Equilibrium time correlation functions in the low-density limit*, *J. Stat. Phys.* **22** (1980), 237–257.
- [9] J. D. BENAMOU and Y. BRENIER, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, *Numer. Math.* **84** (2000), 375–393.
- [10] A. BENSOUSSAN, J.-L. LIONS and G. PAPANICOLAOU, *Boundary layers and homogenization of transport processes*, *Publ. Res. Inst. Math. Sci.* **15** (1979), 53–157.
- [11] L. BERTINI, A. FAGGIONATO and D. GABRIELLI, *Flows, currents, and cycles for Markov chains: large deviation asymptotics*, *Stochastic Process. Appl.* **125** (2015), 2786–2819.
- [12] L. BERTINI, A. DE SOLE, D. GABRIELLI, G. JONA-LASINIO and C. LANDIM, *Large deviations of the empirical current in interacting particle systems*, *Theory Probab. Appl.* **51** (2007), 2–27.
- [13] T. BODINEAU, I. GALLAGHER and L. SAINT-RAYMOND, *The Brownian motion as the limit of a deterministic system of hard-spheres*, *Invent. Math.* **203** (2016), 493–553.
- [14] D. A. DAWSON and J. GÄRTNER, *Large deviations from the McKean-Vlasov limit for weakly interacting diffusions*, *Stochastics* **20** (1987), 247–308.
- [15] P. DEGOND, T. GOUDON and F. POUPAUD, *Diffusion limit for non homogeneous and non-micro-reversible processes*, *Indiana Univ. Math. J.* **49** (2000), 1175–1198.
- [16] M. D. DONSKER and S. R. S. VARADHAN, *Asymptotic evaluation of certain Markov process expectations for large time. I, II*, *Comm. Pure Appl. Math.* **28** (1975), 1–47, 279–301.
- [17] M. ERBAR, *Gradient flows of the entropy for jump processes*, *Ann. Inst. H. Poincaré Probab. Stat.* **50** (2014), 920–945.
- [18] M. ERBAR, *A gradient flow approach to the Boltzmann equation*, arXiv:1603.00540v2, 2017.
- [19] R. ESPOSITO and M. PULVIRENTI, “From Particles to Fluids”, *Hand-Book of Mathematical Fluid Dynamics III*, North-Holland, Amsterdam, 2004.
- [20] G. GALLAVOTTI, “Statistical Mechanics”, Springer, Berlin, 1999.
- [21] N. GIGLI, *On the Heat flow on metric measure spaces: existence, uniqueness and stability*, *Calc. Var. Partial Differential Equations* **39** (2010), 101–120.
- [22] M. JARA, T. KOMOROWSKI and S. OLLA, *A limit theorem for an additive functionals of Markov chains*, *Ann. Appl. Probab.* **19** (2009), 2270–2300.
- [23] C. KIPNIS and S. OLLA, *Large deviations from the hydrodynamical limit for a system of independent Brownian particles*, *Stochastics* **33** (1990), 17–25.
- [24] T. KOMOROWSKI, C. LANDIM and S. OLLA, “Fluctuations in Markov Processes. Time Symmetry and Martingale Approximation”, *Grundlehren der Mathematischen Wissenschaften*, Vol. 345, Springer-Verlag, Berlin, New York, 2012.

- [25] E. LARSEN and J. B. KELLER, *Asymptotic solution of neutron transport problems for small mean free paths*, J. Math. Phys. **15** (1974), 75–81.
- [26] J. L. LEBOWITZ and H. SPOHN, *Steady State Self-Diffusion at Low Density* J. Stat. Phys. **29** (1982), 39–55.
- [27] H. A. LORENTZ, *The motion of electrons in metallic bodies*, Proc. Acad. Amst. **7** (1905), 438–453.
- [28] J. MAAS, *Gradient flows of the entropy for finite Markov chains*, J. Funct. Anal. **261** (2011), 2250–2292.
- [29] A. MELLET, S. MISCHLER and C. MOUHOT, *Fractional diffusion limit for collisional kinetic equations*, Arch. Ration. Mech. Anal. **199** (2011), 493–525.
- [30] A. MIELKE, *Geodesic convexity of the relative entropy in reversible Markov chains*, Calc. Var. Partial Differential Equations **48** (2013), 1–31.
- [31] A. MIELKE, M. A. PELETIER and D. R. M. RENGER, *On the relation between gradient flows and the large-deviation principle, with applications to Markov chains and diffusion*, Potential Anal. **41** (2014), 1293–1327.
- [32] S. MEYN and R. L. TWEEDIE, “Markov Chains and Stochastic Stability”, Second edition, Cambridge University Press, Cambridge, 2009.
- [33] E. SANDIER and S. SERFATY, *Gamma-convergence of gradient flows with applications to Ginzburg-Landau*, Comm. Pure Appl. Math. **57** (2004), 1627–1672.
- [34] S. SERFATY, *Gamma-convergence of gradient flows on Hilbert and metric spaces and applications*, Discrete Contin. Dyn. Syst. **31** (2011), 1427–1451.
- [35] H. SPOHN, *Kinetic equations from Hamiltonian dynamics: Markovian limits*, Rev. Mod. Phys. **52** (1980), 569–615.

Dipartimento di Matematica
“Sapienza” Università di Roma
Piazzale Aldo Moro, 2
00185 Roma, Italia
basile@mat.uniroma1.it
benedetto@mat.uniroma1.it
bertini@mat.uniroma1.it