

# PURE and APPLIED ANALYSIS

# PAM

EMERIC BOUIN, JEAN DOLBEAULT, STÉPHANE MISCHLER,  
CLÉMENT MOUHOT AND CHRISTIAN SCHMEISER

**HYPOCOERCIVITY WITHOUT CONFINEMENT**



vol. 2 no. 2 2020



## HYPOCOERCIVITY WITHOUT CONFINEMENT

EMERIC BOUIN, JEAN DOLBEAULT, STÉPHANE MISCHLER,  
 CLÉMENT MOUHOT AND CHRISTIAN SCHMEISER

Hypocoercivity methods are applied to linear kinetic equations with mass conservation and without confinement in order to prove that the solutions have an algebraic decay rate in the long-time range, which the same as the rate of the heat equation. Two alternative approaches are developed: an analysis based on decoupled Fourier modes and a direct approach where, instead of the Poincaré inequality for the Dirichlet form, Nash’s inequality is employed. The first approach is also used to provide a simple proof of exponential decay to equilibrium on the flat torus. The results are obtained on a space with exponential weights and then extended to larger function spaces by a factorization method. The optimality of the rates is discussed. Algebraic rates of decay on the whole space are improved when the initial datum has moment cancellations.

### 1. Introduction

We consider the Cauchy problem

$$\partial_t f + v \cdot \nabla_x f = \mathbb{L}f, \quad f(0, x, v) = f_0(x, v) \tag{1}$$

for a distribution function  $f(t, x, v)$ , with *position* variable  $x \in \mathbb{R}^d$ , *velocity* variable  $v \in \mathbb{R}^d$ , and with *time*  $t \geq 0$ . Concerning the *collision operator*  $\mathbb{L}$ , we shall consider two cases:

(a) *Fokker–Planck* collision operator:

$$\mathbb{L}f = \nabla_v \cdot [M \nabla_v (M^{-1} f)].$$

(b) *Scattering* collision operator:

$$\mathbb{L}f = \int_{\mathbb{R}^d} \sigma(\cdot, v') (f(v')M(\cdot) - f(\cdot)M(v')) dv'.$$

We shall make the following assumptions on the *local equilibrium*  $M(v)$  and on the *scattering rate*  $\sigma(v, v')$ :

$$\int_{\mathbb{R}^d} M(v) dv = 1, \quad \nabla_v \sqrt{M} \in L^2(\mathbb{R}^d), \quad M \in C(\mathbb{R}^d),$$

$$M = M(|v|), \quad 0 < M(v) \leq c_1 e^{-c_2|v|} \quad \text{for all } v \in \mathbb{R}^d, \quad \text{for some } c_1, c_2 > 0. \tag{H1}$$

*MSC2010:* primary 82C40; secondary 76P05, 35H10, 35K65, 35P15, 35Q84.

*Keywords:* hypocoercivity, linear kinetic equations, Fokker–Planck operator, scattering operator, transport operator, Fourier mode decomposition, Nash’s inequality, factorization method, Green’s function, micro-/macrodecomposition, diffusion limit.

$$1 \leq \sigma(v, v') \leq \bar{\sigma} \quad \text{for all } v, v' \in \mathbb{R}^d, \quad \text{for some } \bar{\sigma} \geq 1. \quad (\text{H2})$$

$$\int_{\mathbb{R}^d} (\sigma(v, v') - \sigma(v', v)) M(v') dv' = 0 \quad \text{for all } v \in \mathbb{R}^d. \quad (\text{H3})$$

Before stating our main results, let us list some preliminary observations.

(i) A typical example of a *local equilibrium* satisfying (H1) is the Gaussian

$$M(v) = \frac{e^{-\frac{|v|^2}{2}}}{(2\pi)^{\frac{d}{2}}}. \quad (2)$$

(ii) With  $\sigma \equiv 1$ , Case (b) includes the relaxation operator  $Lf = M\rho_f - f$ , also known as the *linear BGK operator*, with position density defined by

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

(iii) Positivity and exponential decay of the local equilibrium are essential for our approach. The assumption on the gradient and continuity are technical and only needed for some of our results. Rotational symmetry is not important, but assumed for computational convenience. However the property

$$\int_{\mathbb{R}^d} v M(v) dv = 0,$$

i.e., *zero flux in local equilibrium*, is essential.

(iv) Since microreversibility (or detailed balance), i.e., symmetry of  $\sigma$ , is not required, assumption (H3) is needed for *mass conservation*, i.e.,

$$\int_{\mathbb{R}^d} Lf dv = 0,$$

in Case (b). The boundedness away from zero of  $\sigma$  in (H2) guarantees coercivity of  $L$  relative to its null space (such bound can always be written  $\sigma \geq 1$  by scaling).

Since  $e^{tL}$  propagates probability densities, i.e., conserves mass and nonnegativity,  $L$  dissipates convex relative entropies, implying in particular

$$\int_{\mathbb{R}^d} Lf \frac{f}{M} dv \leq 0.$$

This suggests to use the  $L^2$ -space with the measure  $d\gamma_\infty := \gamma_\infty dv$ , where  $\gamma_\infty(v) = M(v)^{-1}$ , as a functional-analytic framework (the subscript  $\infty$  will make sense later). We shall need the *microscopic coercivity* property

$$-\int_{\mathbb{R}^d} f Lf d\gamma_\infty \geq \lambda_m \int_{\mathbb{R}^d} (f - M\rho_f)^2 d\gamma_\infty, \quad (\text{H4})$$

with some  $\lambda_m > 0$ . In Case (a) it is equivalent to the Poincaré inequality with weight  $M$ ,

$$\int_{\mathbb{R}^d} |\nabla_v h|^2 M dv \geq \lambda_m \int_{\mathbb{R}^d} \left( h - \int_{\mathbb{R}^d} h M dv \right)^2 M dv$$

for all  $h = f/M \in H^1(M dv)$ . It holds as a consequence of the exponential decay assumption in (H1); see, e.g., [Nash 1958; Bakry et al. 2008]. For the normalized Gaussian (2) the optimal constant is known to be  $\lambda_m = 1$ ; see for instance [Beckner 1989]. In Case (b), (H4) means

$$\frac{1}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} \sigma(v, v') M(v) M(v') (u(v) - u(v'))^2 dv' dv \geq \lambda_m \int_{\mathbb{R}^d} (u - \rho_u M)^2 M dv$$

for all  $u = f/M \in L^2(M dv)$ , and it holds with  $\lambda_m = 1$  as a consequence of the lower bound for  $\sigma$  in assumption (H2).

Although the transport operator does not contribute to entropy dissipation, its dispersion in the  $x$ -direction in combination with the dissipative properties of the collision operator yields the desired decay results. In order to perform a *mode-by-mode hypocoercivity* analysis, we introduce the Fourier representation with respect to  $x$ ,

$$f(t, x, v) = \int_{\mathbb{R}^d} \hat{f}(t, \xi, v) e^{+ix \cdot \xi} d\mu(\xi),$$

where  $d\mu(\xi) = (2\pi)^{-d} d\xi$  and  $d\xi$  is the Lebesgue measure on  $\mathbb{R}^d$ . The normalization of  $d\mu(\xi)$  is chosen such that Plancherel's formula reads

$$\|f(t, \cdot, v)\|_{L^2(dx)} = \|\hat{f}(t, \cdot, v)\|_{L^2(d\mu(\xi))},$$

with a straightforward abuse of notation. The Cauchy problem (1) in Fourier variables is now decoupled in the  $\xi$ -direction:

$$\partial_t \hat{f} + i(v \cdot \xi) \hat{f} = \mathbb{L} \hat{f}, \quad \hat{f}(0, \xi, v) = \hat{f}_0(\xi, v). \quad (3)$$

Our main results are devoted to *hypocoercivity without confinement*: when the variable  $x$  is taken in  $\mathbb{R}^d$ , we assume that there is no potential preventing the runaway corresponding to  $|x| \rightarrow +\infty$ . So far, hypocoercivity results have been obtained either in the compact case corresponding to a bounded domain in  $x$ , for instance  $\mathbb{T}^d$ , or in the whole Euclidean space with an external potential  $V$  such that the measure  $e^{-V} dx$  admits a Poincaré inequality. Usually other technical assumptions are required on  $V$  and there are many variants (for instance one can assume a stronger logarithmic Sobolev inequality instead of a Poincaré inequality), but the common property is that some growth condition on  $V$  is assumed and in particular the measure  $e^{-V} dx$  is bounded. Here we consider the case  $V \equiv 0$ , which is obviously a different regime. By replacing the Poincaré inequality by Nash's inequality or using direct estimates in Fourier variables, we adapt the  $L^2$  hypocoercivity methods and prove that an appropriate norm of the solution decays at a rate which is the rate of the heat equation. This observation is compatible with diffusion limits, which have been a source of inspiration for building Lyapunov functionals and establishing the  $L^2$  hypocoercivity method of [Dolbeault et al. 2015]. Before stating any results, we need some notation to implement the *factorization* method of [Gualdani et al. 2017] and obtain estimates in large functional spaces.

Let us consider the measures

$$d\gamma_k := \gamma_k(v) dv, \quad \text{where } \gamma_k(v) = (1 + |v|^2)^{\frac{k}{2}} \quad \text{and } k > d, \quad (4)$$

such that  $1/\gamma_k \in L^1(\mathbb{R}^d)$ . The condition  $k \in (d, \infty]$  then covers the case of weights with a growth of the order of  $|v|^k$ , when  $k$  is finite, and we denote by  $k = \infty$  the case when the weight  $\gamma_\infty = M^{-1}$  grows at least exponentially fast.

**Theorem 1.** *Assume (H1)–(H4),  $x \in \mathbb{R}^d$ , and  $k \in (d, \infty]$ . Then there exists a constant  $C > 0$  such that solutions  $f$  of (1) with initial datum  $f_0 \in L^2(dx d\gamma_k) \cap L^2(d\gamma_k; L^1(dx))$  satisfy, for all  $t \geq 0$ ,*

$$\|f(t, \cdot, \cdot)\|_{L^2(dx d\gamma_k)}^2 \leq C \frac{\|f_0\|_{L^2(dx d\gamma_k)}^2 + \|f_0\|_{L^2(d\gamma_k; L^1(dx))}^2}{(1+t)^{\frac{d}{2}}}.$$

For the heat equation improved decay rates can be shown by Fourier techniques, if the modes with slowest decay are eliminated from the initial data. The following two results are in this spirit.

**Theorem 2.** *Let the assumptions of Theorem 1 hold, and let*

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} f_0 dx dv = 0.$$

*Then there exists  $C > 0$  such that solutions  $f$  of (1) with initial datum  $f_0$  satisfy, for all  $t \geq 0$ ,*

$$\|f(t, \cdot, \cdot)\|_{L^2(dx d\gamma_k)}^2 \leq C \frac{\|f_0\|_{L^2(d\gamma_{k+2}; L^1(dx))}^2 + \|f_0\|_{L^2(d\gamma_k; L^1(|x| dx))}^2 + \|f_0\|_{L^2(dx d\gamma_k)}^2}{(1+t)^{\frac{d}{2}+1}},$$

*with  $k \in (d, \infty)$ .*

The case of Theorem 2, but with  $k = \infty$ , is covered in Theorem 3 under the stronger assumption that  $M$  is a Gaussian. For the formulation of a result corresponding to the cancellation of higher-order moments, we introduce the set  $\mathbb{R}_\ell[X, V]$  of polynomials of order at most  $\ell$  in the variables  $X, V \in \mathbb{R}^d$  (the sum of the degrees in  $X$  and in  $V$  is at most  $\ell$ ). We also need that the kernel of the collision operator is spanned by a Gaussian function in order to keep polynomial spaces invariant. This means that for any  $P \in \mathbb{R}_\ell[X, V]$ , one has  $(L - T)(PM) \in \mathbb{R}_\ell[X, V]M$ . Since the transport operator mixes both variables  $x$  and  $v$ , one needs moments with respect to both  $x$  and  $v$  variables.

**Theorem 3.** *In Case (a), let  $M$  be the normalized Gaussian (2). In Case (b), we assume that  $\sigma \equiv 1$ . Let  $k \in (d, \infty]$ ,  $\ell \in \mathbb{N}$ , and assume that the initial datum  $f_0 \in L^1(\mathbb{R}^d \times \mathbb{R}^d)$  is such that*

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} f_0(x, v) P(x, v) dx dv = 0 \tag{5}$$

*for all  $P \in \mathbb{R}_\ell[X, V]$ . Then there exists a constant  $c_k > 0$  such that any solution  $f$  of (1) with initial datum  $f_0$  satisfies, for all  $t \geq 0$ ,*

$$\|f(t, \cdot, \cdot)\|_{L^2(dx d\gamma_k)}^2 \leq c_k \frac{\|f_0\|_{L^2(d\gamma_{k+2}; L^1(dx))}^2 + \|f_0\|_{L^2(d\gamma_k; L^1(|x| dx))}^2 + \|f_0\|_{L^2(dx d\gamma_k)}^2}{(1+t)^{\frac{d}{2}+1+\ell}}.$$

The outline of this paper goes as follows. In Section 2, we slightly strengthen the abstract hypocoercivity result of [Dolbeault et al. 2015] by allowing complex Hilbert spaces and by providing explicit formulas for the coefficients in the decay rate (Proposition 4). In Corollary 5, this result is applied for fixed  $\xi$  to

the Fourier transformed problem (3), where integrals are computed with respect to the measure  $d\gamma_\infty$  in the velocity variable  $v$ . Since the frequency  $\xi$  can be considered as a parameter, we shall speak of a *mode-by-mode hypocoercivity* result. It provides exponential decay, however, with a rate deteriorating as  $\xi \rightarrow 0$ .

In Section 3, we state a special case (Proposition 6) of the *factorization* result of [Gualdani et al. 2017] with explicit constants which corresponds to an *enlargement* of the space, and also a *shrinking* result (Proposition 7) which will be useful in Section 6.2. By the enlargement result, the estimate corresponding to the exponential weight  $\gamma_\infty$  is extended in Corollary 8 to larger spaces corresponding to the algebraic weights  $\gamma_k$  with  $k \in (d, \infty)$ . As a straightforward consequence, in Section 4, we recover an *exponential convergence rate* in the case of the flat torus  $\mathbb{T}^d$  (Corollary 9), and then give a first proof of the *algebraic decay rate* of Theorem 1 in the whole space without confinement.

In Section 5, a hypocoercivity method, where the Poincaré inequality, or the so-called *macroscopic coercivity* condition, is replaced by the *Nash inequality*, provides an alternative proof of Theorem 1. Such a direct approach is also applicable to problems with nonconstant coefficients like scattering operators with  $x$ -dependent scattering rates  $\sigma$ , or Fokker–Planck operators with  $x$ -dependent diffusion constants like  $\nabla_v \cdot (\mathcal{D}(x)M\nabla_v(M^{-1}f))$ .

The *improved algebraic decay rates* of Theorems 2 and 3 are obtained by direct Fourier estimates in Section 6. As we shall see in Appendix A, the rates of Theorem 1 are optimal: the decay rate is the rate of the heat equation on  $\mathbb{R}^d$ . Our method is consistent with the *diffusion limit* and provides estimates which are asymptotically uniform in this regime: see Appendix B. We also check that the results of Theorems 2 and 3 are uniform in the diffusive limit in Appendix B.

We conclude this introduction by a brief *review of the literature*: On the whole Euclidean space, we refer to [Vázquez 2017] for recent lecture notes on available techniques for capturing the large-time asymptotics of the heat equation. Some of our results make a clear link with the heat flow seen as the diffusion limit of the kinetic equation. We also refer to [Iacobucci et al. 2019] for recent results on the diffusion limit, or overdamped limit (see Appendix B).

The mode-by-mode analysis is an extension of the hypocoercivity theory of [Dolbeault et al. 2015], which has been inspired by [Hérau 2006], but is also close to the Kawashima compensating function method [1990]; see also [Glassey 1996, Chapter 3, Section 3.9]. We also refer to [Duan 2011] where the Kawashima approach is applied to the Fokker–Planck operator (a) and to a particular case of the scattering model (b).

The word *hypocoercivity* was coined by T. Gallay and widely disseminated in the context of kinetic theory by C. Villani. In [Mouhot and Neumann 2006; Villani 2006; 2009], the method deals with large-time properties of the solutions by considering an  $H^1$ -norm (in  $x$  and  $v$  variables) and taking into account cross-terms. This is very well explained in [Villani 2006, Section 3], but was already present in earlier works like [Hérau and Nier 2004]. Hypocoercivity theory is inspired by and related to the earlier *hypoellipticity* theory. The latter has a long history in the context of the kinetic Fokker–Planck equation. One can refer for instance to [Eckmann and Hairer 2003; Hérau and Nier 2004] and much earlier to Hörmander’s theory [1967]. The seed for such an approach can even be traced back to Kolmogorov’s computation [1934] of Green’s kernel for the kinetic Fokker–Planck equation, which has

been reconsidered in [Ilin and Hasminskii 1964] and successfully applied, for instance, to the study of the Vlasov–Poisson–Fokker–Planck system in [Victory and O’Dwyer 1990; Bouchut 1993].

Linear Boltzmann equations and BGK (Bhatnagar–Gross–Krook, see [Bhatnagar et al. 1954]) models also have a long history: we refer to [Degond et al. 2000; Cáceres et al. 2003] for key mathematical properties, and to [Mouhot and Neumann 2006; Hérau 2006] for first hypocoercivity results. In this paper we will mostly rely on [Dolbeault et al. 2009; 2015]. However, among more recent contributions, one has to quote [Han-Kwan and Léautaud 2015; Achleitner et al. 2016; Bouin et al. 2017] and also an approach based on the Fisher information which has recently been implemented in [Evans 2017; Monmarché 2017].

With the *exponential weight*  $\gamma_\infty = M^{-1}$ , Corollary 9 can be obtained directly by the method of [Dolbeault et al. 2015]. In this paper we also obtain a result for weights with polynomial growth in the velocity variable based on [Gualdani et al. 2017]. For completeness, let us mention that recently the exponential growth issue was overcome for the Fokker–Planck case in [Kavian and Mischler 2015; Mischler and Mouhot 2016] by a different method. The improved decay rates established in Theorems 2 and 3 generalize to kinetic models similar results known for the heat equation; see for instance [Mischler and Mouhot 2016, Remark 3.2(7)] or [Bartier et al. 2011].

## 2. Mode-by-mode hypocoercivity

Let us consider the evolution equation

$$\frac{dF}{dt} + \mathbb{T}F = \mathbb{L}F, \quad (6)$$

where  $\mathbb{T}$  and  $\mathbb{L}$  are respectively a general *transport operator* and a general *linear collision operator*. We shall use the abstract approach of [Dolbeault et al. 2015]. Although the extension of the method to Hilbert spaces over complex numbers is rather straightforward, we carry it out here for completeness. For details on the Cauchy problem or, e.g., on the domains of the operators, we refer to [Dolbeault et al. 2015]. Notice that we do not ask that  $\mathbb{L}$  is a Hermitian operator but simply assume that  $\mathbb{L}^*A = 0$ .

**Proposition 4.** *Let  $\mathbb{L}$  and  $\mathbb{T}$  be closed unbounded linear operators on the complex Hilbert space  $(\mathcal{H}, \langle \cdot, \cdot \rangle)$  with dense domains  $\mathcal{D}(\mathbb{L})$  and  $\mathcal{D}(\mathbb{T})$ . Assume that  $\mathbb{T}$  is anti-Hermitian. Let  $\Pi$  be the orthogonal projection onto the null space of  $\mathbb{L}$  and define*

$$A := (1 + (\mathbb{T}\Pi)^*\mathbb{T}\Pi)^{-1}(\mathbb{T}\Pi)^*,$$

where  $*$  denotes the adjoint with respect to  $\langle \cdot, \cdot \rangle$ . We assume that  $\mathbb{L}^*A = 0$  and that there are positive constants  $\lambda_m, \lambda_M$ , and  $C_M$  such that, for any  $F \in \mathcal{H}$ , the following properties hold:

- Microscopic coercivity:

$$-\langle \mathbb{L}F, F \rangle \geq \lambda_m \|(1 - \Pi)F\|^2 \quad \text{for all } F \in \mathcal{D}(\mathbb{L}). \quad (\text{A1})$$

- Macroscopic coercivity:

$$\|\mathbb{T}\Pi F\|^2 \geq \lambda_M \|\Pi F\|^2 \quad \text{for all } F \in \mathcal{D}(\mathbb{T}). \quad (\text{A2})$$

- Parabolic macroscopic dynamics:

$$\Pi T \Pi F = 0 \quad \text{for all } F \in \mathcal{D}(T). \quad (\text{A3})$$

- Bounded auxiliary operators:

$$\|\text{AT}(1 - \Pi)F\| + \|\text{AL}F\| \leq C_M \|(1 - \Pi)F\| \quad \text{for all } F \in \mathcal{D}(L) \cap \mathcal{D}(T). \quad (\text{A4})$$

Then  $L - T$  generates a  $C_0$ -semigroup and for any  $t \geq 0$  we have

$$\|e^{(L-T)t}\|^2 \leq 3e^{-\lambda t}, \quad \text{where } \lambda = \frac{\lambda_M}{3(1 + \lambda_M)} \min \left\{ 1, \lambda_m, \frac{\lambda_m \lambda_M}{(1 + \lambda_M)C_M^2} \right\}. \quad (7)$$

*Proof.* For some  $\delta > 0$  to be determined later, the Lyapunov functional

$$\text{H}[F] := \frac{1}{2}\|F\|^2 + \delta \text{Re}\langle \text{AF}, F \rangle$$

is such that  $\frac{d}{dt}\text{H}[F] = -\text{D}[F]$  if  $F$  solves (6), with

$$\text{D}[F] := -\langle \text{LF}, F \rangle + \delta \langle \text{AT}\Pi F, F \rangle + \delta \text{Re}\langle \text{AT}(1 - \Pi)F, F \rangle - \delta \text{Re}\langle \text{TAF}, F \rangle - \delta \text{Re}\langle \text{AL}F, F \rangle.$$

Note that we have used the fact that  $\text{Re}\langle \text{AF}, \text{LF} \rangle = 0$  because of the assumption  $\text{L}^*A = 0$ , and also that  $\langle \text{AT}\Pi F, F \rangle$  is real because  $\text{AT}\Pi$  is self-adjoint by construction. Since the Hermitian operator  $\text{AT}\Pi$  can be interpreted as the application of the map  $z \mapsto (1 + z)^{-1}z$  to  $(\text{T}\Pi)^*\text{T}\Pi$  and as a consequence of the spectral theorem [Reed and Simon 1980, Theorem VII.2, p. 225], the conditions (A1) and (A2) imply

$$-\langle \text{LF}, F \rangle + \delta \langle \text{AT}\Pi F, F \rangle \geq \lambda_m \|(1 - \Pi)F\|^2 + \frac{\delta \lambda_M}{1 + \lambda_M} \|\Pi F\|^2.$$

As in [Dolbeault et al. 2015, Lemma 1], if  $G = \text{AF}$ , i.e.,  $G + (\text{T}\Pi)^*\text{T}\Pi G = (\text{T}\Pi)^*F$ , one has

$$\|\text{AF}\|^2 + \|\text{TAF}\|^2 = \langle G, G + (\text{T}\Pi)^*\text{T}\Pi G \rangle = \langle G, (\text{T}\Pi)^*F \rangle = \langle \text{TAF}, (1 - \Pi)F \rangle,$$

where we have used  $A = \Pi A$  and  $\Pi T \Pi = 0$ . Using  $|\langle \text{TAF}, (1 - \Pi)F \rangle| \leq \|\text{TAF}\|^2 + \frac{1}{4}\|(1 - \Pi)F\|^2$ , one gets

$$\|\text{AF}\|^2 \leq \frac{1}{4}\|(1 - \Pi)F\|^2, \quad (8)$$

which implies that  $|\text{Re}\langle \text{AF}, F \rangle| \leq \|\text{AF}\|\|F\| \leq \frac{1}{2}\|F\|^2$  and provides us with the norm equivalence of  $\text{H}[F]$  and  $\|F\|^2$ ,

$$\frac{1}{2}(1 - \delta)\|F\|^2 \leq \text{H}[F] \leq \frac{1}{2}(1 + \delta)\|F\|^2. \quad (9)$$

With  $X := \|(1 - \Pi)F\|$  and  $Y := \|\Pi F\|$ , it follows from (A4) that

$$\text{D}[F] \geq (\lambda_m - \delta)X^2 + \frac{\delta \lambda_M}{1 + \lambda_M} Y^2 - \delta C_M XY.$$

The choice

$$\delta = \frac{1}{2} \min \left\{ 1, \lambda_m, \frac{\lambda_m \lambda_M}{(1 + \lambda_M)C_M^2} \right\}$$

implies

$$D[F] \geq \frac{\lambda_m}{4} X^2 + \frac{\delta \lambda_M}{2(1+\lambda_M)} Y^2 \geq \frac{1}{4} \min \left\{ \lambda_m, \frac{2\delta \lambda_M}{1+\lambda_M} \right\} \|F\|^2 \geq \frac{2\delta \lambda_M}{3(1+\lambda_M)} \mathsf{H}[F].$$

With  $\lambda$  defined in (7), using  $\delta \leq \frac{1}{2}$  and  $(1+\delta)/(1-\delta) \leq 3$ , we get

$$\|F(t)\|^2 \leq \frac{2}{1-\delta} \mathsf{H}[F](t) \leq \frac{1+\delta}{1-\delta} e^{-\lambda t} \|F(0)\|^2 \leq 3e^{-\lambda t} \|F(0)\|^2. \quad \square$$

For any fixed  $\xi \in \mathbb{R}^d$ , let us apply Proposition 4 to (3) with  $F = \hat{f}$  and

$$\mathcal{H} = L^2(d\gamma_\infty), \quad \|F\|^2 = \int_{\mathbb{R}^d} |F|^2 d\gamma_\infty, \quad \Pi F = M \int_{\mathbb{R}^d} F dv = M\rho_F, \quad \Upsilon F = i(v \cdot \xi)F.$$

Here we are in a mode-by-mode framework in which the transport operator  $\Upsilon$  is a simple multiplication operator.

**Corollary 5.** *Assume (H1)–(H4), and take  $\xi \in \mathbb{R}^d$ . If  $\hat{f}$  is a solution of (3) such that  $\hat{f}_0(\xi, \cdot) \in L^2(d\gamma_\infty)$ , then for any  $t \geq 0$  we have*

$$\|\hat{f}(t, \xi, \cdot)\|_{L^2(d\gamma_\infty)}^2 \leq 3e^{-\mu_\xi t} \|\hat{f}_0(\xi, \cdot)\|_{L^2(d\gamma_\infty)}^2,$$

where

$$\mu_\xi := \frac{\Lambda |\xi|^2}{1 + |\xi|^2} \quad \text{and} \quad \Lambda = \frac{1}{3} \min\{1, \Theta\} \min\left\{1, \frac{\lambda_m \Theta^2}{K + \Theta \kappa^2}\right\}, \quad (10)$$

with

$$\Theta := \int_{\mathbb{R}^d} (v \cdot e)^2 M(v) dv, \quad K := \int_{\mathbb{R}^d} (v \cdot e)^4 M(v) dv, \quad \theta := \frac{4}{d} \int_{\mathbb{R}^d} |\nabla_v \sqrt{M}|^2 dv \quad (11)$$

for an arbitrary  $e \in \mathbb{S}^{d-1}$ , and with  $\kappa = \sqrt{\theta}$  in Case (a) and  $\kappa = 2\bar{\sigma} \sqrt{\Theta}$  in Case (b).

*Proof.* We check that the assumptions of Proposition 4 are satisfied with  $F = \hat{f}$ . The property  $L^*A = 0$  is a consequence of the mass conservation  $\int_{\mathbb{R}^d} Lf dv = 0$  because  $\Pi A = A$ . Assumption (H4) implies (A1). Concerning the macroscopic coercivity (A2), since

$$\Upsilon \Pi F = i(v \cdot \xi) \rho_F M,$$

one has

$$\|\Upsilon \Pi F\|^2 = |\rho_F|^2 \int_{\mathbb{R}^d} |v \cdot \xi|^2 M(v) dv = \Theta |\xi|^2 |\rho_F|^2 = \Theta |\xi|^2 \|\Pi F\|^2,$$

and thus (A2) holds with  $\lambda_M = \Theta |\xi|^2$ . By assumption  $M(v)$  depends only on  $|v|$ , so it is *unbiased*, i.e.,  $\int_{\mathbb{R}^d} v M(v) dv = 0$ , which means that (A3) holds.

Let us now prove (A4). Since  $(\Upsilon \Pi)^* F = -\Pi \Upsilon F = -i(\xi \int_{\mathbb{R}^d} v' F(v') dv') M$ , we obtain

$$(1 + (\Upsilon \Pi)^* \Upsilon \Pi) \rho M = \left(1 + \int_{\mathbb{R}^d} (\xi \cdot v')^2 M(v') dv'\right) \rho M = (1 + \Theta |\xi|^2) \rho M$$

and the operator  $A$ , defined in Proposition 4, is given mode-by-mode by

$$AF = \frac{-i\xi \int_{\mathbb{R}^d} v' F(v') dv'}{1 + \Theta|\xi|^2} M.$$

As a consequence,  $A$  satisfies the estimate

$$\begin{aligned} \|AF\| &= \|A(1 - \Pi)F\| \leq \frac{1}{1 + \Theta|\xi|^2} \int_{\mathbb{R}^d} \frac{|(1 - \Pi)F|}{\sqrt{M}} |v \cdot \xi| \sqrt{M} dv \\ &\leq \frac{\|(1 - \Pi)F\|}{1 + \Theta|\xi|^2} \left( \int_{\mathbb{R}^d} (v \cdot \xi)^2 M dv \right)^{\frac{1}{2}} = \frac{\sqrt{\Theta}|\xi|}{1 + \Theta|\xi|^2} \|(1 - \Pi)F\|. \end{aligned}$$

In Case (b) the collision operator  $L$  is obviously bounded,

$$\|LF\| \leq 2\bar{\sigma} \|(1 - \Pi)F\|$$

and, as a consequence,

$$\|ALF\| \leq \frac{2\bar{\sigma}\sqrt{\Theta}|\xi|}{1 + \Theta|\xi|^2} \|(1 - \Pi)F\|.$$

We also notice that  $L^*A = 0$  according to (H3). For estimating  $AL$  in Case (a), we note that

$$\int_{\mathbb{R}^d} v LF dv = 2 \int_{\mathbb{R}^d} \nabla_v \sqrt{M} \frac{F}{\sqrt{M}} dv$$

and obtain as above that

$$\|ALF\| \leq \frac{2}{1 + \Theta|\xi|^2} \int_{\mathbb{R}^d} \frac{|(1 - \Pi)F|}{\sqrt{M}} |\xi \cdot \nabla_v \sqrt{M}| dv \leq \frac{\sqrt{\Theta}|\xi|}{1 + \Theta|\xi|^2} \|(1 - \Pi)F\|.$$

For both cases we finally obtain

$$\|ALF\| \leq \frac{\kappa|\xi|}{1 + \Theta|\xi|^2} \|(1 - \Pi)F\|.$$

Similarly we can estimate

$$AT(1 - \Pi)F = \frac{\int_{\mathbb{R}^d} (v' \cdot \xi)^2 (1 - \Pi)F(v') dv'}{1 + \Theta|\xi|^2} M$$

by

$$\begin{aligned} \|AT(1 - \Pi)F\| &= \frac{|\int_{\mathbb{R}^d} (v' \cdot \xi)^2 (1 - \Pi)F(v') dv'|}{1 + \Theta|\xi|^2} \\ &\leq \frac{(\int_{\mathbb{R}^d} (v' \cdot \xi)^4 M(v') dv')^{\frac{1}{2}}}{1 + \Theta|\xi|^2} \|(1 - \Pi)F\| = \frac{\sqrt{K}|\xi|^2}{1 + \Theta|\xi|^2} \|(1 - \Pi)F\|, \end{aligned}$$

meaning that we have proven (A4) with

$$C_M = \frac{\kappa|\xi| + \sqrt{K}|\xi|^2}{1 + \Theta|\xi|^2}.$$

With the elementary estimates

$$\frac{\Theta|\xi|^2}{1+\Theta|\xi|^2} \geq \min\{1, \Theta\} \frac{|\xi|^2}{1+|\xi|^2} \quad \text{and} \quad \frac{\lambda_M}{(1+\lambda_M)C_M^2} = \frac{\Theta(1+\Theta|\xi|^2)}{(k+\sqrt{K}|\xi|)^2} \geq \frac{\Theta^2}{K+\Theta k^2},$$

the proof is completed using (7).  $\square$

### 3. Enlarging and shrinking spaces by factorization

Square integrability against the inverse of the *local equilibrium*  $M$  is a rather restrictive assumption on the initial datum. In this section it will be relaxed with the help of the abstract *factorization method* of [Gualdani et al. 2017] in a simple case (factorization of order 1). Here we state the result and sketch a proof in a special case, for the convenience of the reader. We shall then give a result based on similar computations in the opposite direction: how to establish a rate in a stronger norm, which corresponds to a *shrinking* of the functional space. We will conclude with an application to the problem studied in Corollary 5. Let us start by *enlarging* the space.

**Proposition 6.** *Let  $\mathcal{B}_1, \mathcal{B}_2$  be Banach spaces and let  $\mathcal{B}_2$  be continuously imbedded in  $\mathcal{B}_1$ , i.e.,  $\|\cdot\|_1 \leq c_1 \|\cdot\|_2$ . Let  $\mathfrak{B}$  and  $\mathfrak{A} + \mathfrak{B}$  be the generators of the strongly continuous semigroups  $e^{\mathfrak{B}t}$  and  $e^{(\mathfrak{A}+\mathfrak{B})t}$  on  $\mathcal{B}_1$ . Assume that there are positive constants  $c_2, c_3, c_4, \lambda_1$  and  $\lambda_2$  such that, for all  $t \geq 0$ ,*

$$\|e^{(\mathfrak{A}+\mathfrak{B})t}\|_{2 \rightarrow 2} \leq c_2 e^{-\lambda_2 t}, \quad \|e^{\mathfrak{B}t}\|_{1 \rightarrow 1} \leq c_3 e^{-\lambda_1 t}, \quad \|\mathfrak{A}\|_{1 \rightarrow 2} \leq c_4,$$

where  $\|\cdot\|_{i \rightarrow j}$  denotes the operator norm for linear mappings from  $\mathcal{B}_i$  to  $\mathcal{B}_j$ . Then there exists a positive constant  $C = C(c_1, c_2, c_3, c_4)$  such that, for all  $t \geq 0$ ,

$$\|e^{(\mathfrak{A}+\mathfrak{B})t}\|_{1 \rightarrow 1} \leq \begin{cases} C(1+|\lambda_1-\lambda_2|^{-1})e^{-\min\{\lambda_1, \lambda_2\}t} & \text{for } \lambda_1 \neq \lambda_2, \\ C(1+t)e^{-\lambda_1 t} & \text{for } \lambda_1 = \lambda_2. \end{cases}$$

*Proof.* Integrating the identity  $\frac{d}{ds}(e^{(\mathfrak{A}+\mathfrak{B})s}e^{\mathfrak{B}(t-s)}) = e^{(\mathfrak{A}+\mathfrak{B})s}\mathfrak{A}e^{\mathfrak{B}(t-s)}$  with respect to  $s \in [0, t]$  gives

$$e^{(\mathfrak{A}+\mathfrak{B})t} = e^{\mathfrak{B}t} + \int_0^t e^{(\mathfrak{A}+\mathfrak{B})s}\mathfrak{A}e^{\mathfrak{B}(t-s)} ds.$$

The proof is completed by the straightforward computation

$$\begin{aligned} \|e^{(\mathfrak{A}+\mathfrak{B})t}\|_{1 \rightarrow 1} &\leq c_3 e^{-\lambda_1 t} + c_1 \int_0^t \|e^{(\mathfrak{A}+\mathfrak{B})s}\mathfrak{A}e^{\mathfrak{B}(t-s)}\|_{1 \rightarrow 2} ds \\ &\leq c_3 e^{-\lambda_1 t} + c_1 c_2 c_3 c_4 e^{-\lambda_1 t} \int_0^t e^{(\lambda_1 - \lambda_2)s} ds. \end{aligned} \quad \square$$

The second statement of this section is devoted to a result on the *shrinking* of the functional space. It is based on a computation which is similar to the one of the proof of Proposition 6.

**Proposition 7.** *Let  $\mathcal{B}_1, \mathcal{B}_2$  be Banach spaces and let  $\mathcal{B}_2$  be continuously imbedded in  $\mathcal{B}_1$ , i.e.,  $\|\cdot\|_1 \leq c_1 \|\cdot\|_2$ . Let  $\mathfrak{B}$  and  $\mathfrak{A} + \mathfrak{B}$  be the generators of the strongly continuous semigroups  $e^{\mathfrak{B}t}$  and  $e^{(\mathfrak{A}+\mathfrak{B})t}$*

on  $\mathcal{B}_1$ . Assume that there are positive constants  $c_2, c_3, c_4, \lambda_1$  and  $\lambda_2$  such that, for all  $t \geq 0$ ,

$$\|e^{(\mathfrak{A}+\mathfrak{B})t}\|_{1 \rightarrow 1} \leq c_2 e^{-\lambda_1 t}, \quad \|e^{\mathfrak{B}t}\|_{2 \rightarrow 2} \leq c_3 e^{-\lambda_2 t}, \quad \|\mathfrak{A}\|_{1 \rightarrow 2} \leq c_4,$$

where  $\|\cdot\|_{i \rightarrow j}$  denotes the operator norm for linear mappings from  $\mathcal{B}_i$  to  $\mathcal{B}_j$ . Then there exists a positive constant  $C = C(c_1, c_2, c_3, c_4)$  such that, for all  $t \geq 0$ ,

$$\|e^{(\mathfrak{A}+\mathfrak{B})t}\|_{2 \rightarrow 2} \leq \begin{cases} C(1 + |\lambda_2 - \lambda_1|^{-1})e^{-\min\{\lambda_2, \lambda_1\}t} & \text{for } \lambda_2 \neq \lambda_1, \\ C(1+t)e^{-\lambda_1 t} & \text{for } \lambda_1 = \lambda_2. \end{cases}$$

*Proof.* Integrating the identity  $\frac{d}{ds}(e^{\mathfrak{B}(t-s)}e^{(\mathfrak{A}+\mathfrak{B})s}) = e^{\mathfrak{B}(t-s)}\mathfrak{A}e^{(\mathfrak{A}+\mathfrak{B})s}$  with respect to  $s \in [0, t]$  gives

$$e^{(\mathfrak{A}+\mathfrak{B})t} = e^{\mathfrak{B}t} + \int_0^t e^{\mathfrak{B}(t-s)}\mathfrak{A}e^{(\mathfrak{A}+\mathfrak{B})s} ds.$$

The proof is completed by the straightforward computation

$$\begin{aligned} \|e^{(\mathfrak{A}+\mathfrak{B})t}\|_{2 \rightarrow 2} &\leq c_3 e^{-\lambda_2 t} + \int_0^t \|e^{\mathfrak{B}(t-s)}\mathfrak{A}e^{(\mathfrak{A}+\mathfrak{B})s}\|_{2 \rightarrow 2} ds \\ &\leq c_3 e^{-\lambda_2 t} + c_1 \int_0^t \|e^{\mathfrak{B}(t-s)}\mathfrak{A}e^{(\mathfrak{A}+\mathfrak{B})s}\|_{1 \rightarrow 2} ds \\ &\leq c_3 e^{-\lambda_2 t} + c_1 \int_0^t \|e^{\mathfrak{B}(t-s)}\|_{2 \rightarrow 2} \|\mathfrak{A}\|_{1 \rightarrow 2} \|e^{(\mathfrak{A}+\mathfrak{B})s}\|_{1 \rightarrow 1} ds \\ &\leq c_3 e^{-\lambda_2 t} + c_1 c_2 c_3 c_4 e^{-\lambda_2 t} \int_0^t e^{(\lambda_2 - \lambda_1)s} ds. \quad \square \end{aligned}$$

We will use Proposition 7 in Section 6.2. Coming back to the problem studied in Corollary 5, Proposition 6 applies to (3) with the spaces  $\mathcal{B}_1 = L^2(d\gamma_k)$ ,  $k \in (d, \infty)$ , and  $\mathcal{B}_2 = L^2(d\gamma_\infty)$  corresponding to the weights defined by (4). The exponential growth of  $\gamma_\infty$  guarantees that  $\mathcal{B}_2$  is continuously imbedded in  $\mathcal{B}_1$ .

**Corollary 8.** Assume (H1)–(H4),  $k \in (d, \infty]$ , and  $\xi \in \mathbb{R}^d$ . Then there exists a constant  $C > 0$  such that solutions  $\hat{f}$  of (3) with initial datum  $\hat{f}_0(\xi, \cdot) \in L^2(d\gamma_k)$  satisfy, with  $\mu_\xi$  given by (10),

$$\|\hat{f}(t, \xi, \cdot)\|_{L^2(d\gamma_k)}^2 \leq C e^{-\mu_\xi t} \|\hat{f}_0(\xi, \cdot)\|_{L^2(d\gamma_k)}^2 \quad \text{for all } t \geq 0.$$

*Proof.* In Case (a), let us define  $\mathfrak{A}$  and  $\mathfrak{B}$  by  $\mathfrak{A}F = N\chi_R F$  and  $\mathfrak{B}F = -i(v \cdot \xi)F + LF - \mathfrak{A}F$ , where  $N$  and  $R$  are two positive constants,  $\chi$  is a smooth function such that  $\mathbb{1}_{B_1} \leq \chi \leq \mathbb{1}_{B_2}$ , and  $\chi_R := \chi(\cdot/R)$ . Here  $B_r$  is the centered ball of radius  $r$ . It has been established in [Mischler and Mouhot 2016, Lemma 3.8] that if  $k > d$ , then the inequality

$$\int_{\mathbb{R}^d} (\mathbb{L} - \mathfrak{A})(F)F d\gamma_k \leq -\lambda_1 \int_{\mathbb{R}^d} F^2 d\gamma_k$$

holds for some  $\lambda_1 > 0$ . Moreover,  $\lambda_1$  can be chosen arbitrarily large for  $R$  and  $N$  large enough. The boundedness of  $\mathfrak{A} : \mathcal{B}_1 \rightarrow \mathcal{B}_2$  follows from the compactness of the support of  $\chi$  and Proposition 6 applies with  $\lambda_2 = \frac{1}{2}\mu_\xi \leq \frac{1}{4}$ , where  $\mu_\xi$  is given by (10).

In Case (b), we consider  $\mathfrak{A}$  and  $\mathfrak{B}$  such that

$$\begin{aligned}\mathfrak{A}F(v) &= M(v) \int_{\mathbb{R}^d} \sigma(v, v') F(v') dv', \\ \mathfrak{B}F(v) &= - \left[ i(v \cdot \xi) + \int_{\mathbb{R}^d} \sigma(v, v') M(v') dv' \right] F(v).\end{aligned}$$

The boundedness of  $\mathfrak{A} : \mathcal{B}_1 \rightarrow \mathcal{B}_2$  follows from (H2) and

$$\|\mathfrak{A}F\|_{L^2(d\gamma_\infty)} \leq \bar{\sigma} \|F\|_{L^1(dv)} \leq \bar{\sigma} \left( \int_{\mathbb{R}^d} \gamma_k^{-1} dv \right)^{\frac{1}{2}} \|F\|_{L^2(d\gamma_k)}.$$

Proposition 6 applies with  $\lambda_2 = \frac{1}{2}\mu_\xi \leq \frac{1}{4}$  and  $\lambda_1 = 1$  because  $\int_{\mathbb{R}^d} \sigma(v, v') M(v') dv' \geq 1$ .  $\square$

#### 4. Asymptotic behavior based on mode-by-mode estimates

In this section we consider (1) and use the estimates of Corollary 5 with weight  $\gamma_\infty = 1/M$  and Corollary 8 for weights with  $O(|v|^k)$  growth to get decay rates with respect to  $t$ . We shall consider two cases for the spatial variable  $x$ . In Section 4.1, we assume that  $x \in \mathbb{T}^d$ , where  $\mathbb{T}^d$  is the flat  $d$ -dimensional torus (represented by  $[0, 2\pi)^d$  with periodic boundary conditions) and prove an exponential convergence rate. In Section 4.2, we assume that  $x \in \mathbb{R}^d$  and establish algebraic decay rates.

**4.1. Exponential convergence to equilibrium in  $\mathbb{T}^d$ .** In the periodic case  $x \in \mathbb{T}^d$  there is a unique nonzero normalized equilibrium given by

$$f_\infty(x, v) = \rho_\infty M(v) \quad \text{with} \quad \rho_\infty = \frac{1}{|\mathbb{T}^d|} \iint_{\mathbb{T}^d \times \mathbb{R}^d} f_0 dx dv.$$

**Corollary 9.** *Assume (H1)–(H4) and  $k \in (d, \infty]$ . Then there exists a constant  $C > 0$  such that the solution  $f$  of (1) on  $\mathbb{T}^d \times \mathbb{R}^d$  with initial datum  $f_0 \in L^2(dx d\gamma_k)$  satisfies, with  $\Lambda$  given by (10),*

$$\|f(t, \cdot, \cdot) - f_\infty\|_{L^2(dx d\gamma_k)} \leq C \|f_0 - f_\infty\|_{L^2(dx d\gamma_k)} e^{-\Lambda \frac{t}{4}} \quad \text{for all } t \geq 0.$$

*Proof.* We represent the flat torus  $\mathbb{T}^d$  by  $[0, 2\pi)^d$  with periodic boundary conditions, and the Fourier variable is denoted by  $\xi \in \mathbb{Z}^d$ . For  $\xi = 0$ , the microscopic coercivity (see Section 2) implies

$$\|\hat{f}(t, 0, \cdot) - \hat{f}_\infty(0, \cdot)\|_{L^2(d\gamma_\infty)} \leq \|\hat{f}_0(0, \cdot) - \hat{f}_\infty(0, \cdot)\|_{L^2(d\gamma_\infty)} e^{-t}.$$

For all other modes,  $\hat{f}_\infty(\xi, \cdot) = 0$  for any  $\xi \neq 0$  (that is, for any  $\xi$  such that  $|\xi| \geq 1$ ). We can use Corollary 5 with  $\mu_\xi \geq \frac{1}{2}\Lambda$ , with the notation of (10). An application of Parseval's identity then proves the result for  $k = \infty$  and  $C = \sqrt{3}$ . If  $k$  is finite, the result with the weight  $\gamma_k$  follows from Corollary 8.  $\square$

Note that the latter result can also alternatively be proved by directly applying Proposition 4 to (1), as in [Dolbeault et al. 2015].

**4.2. Algebraic decay rates in  $\mathbb{R}^d$ .** With the result of Corollaries 5 and 8 we obtain a first proof of Theorem 1 as follows. Let  $C > 0$  be a generic constant which is going to change from line to line. Plancherel's formula implies

$$\|f(t, \cdot, \cdot)\|_{L^2(dx d\gamma_k)}^2 \leq C \int_{\mathbb{R}^d} \left( \int_{\mathbb{R}^d} e^{-\mu_\xi t} |\hat{f}_0|^2 d\xi \right) d\gamma_k.$$

We know that

$$\int_{|\xi| \leq 1} e^{-\mu_\xi t} d\xi \leq \int_{\mathbb{R}^d} e^{-\frac{\Lambda}{2} |\xi|^2 t} d\xi = \left( \frac{2\pi}{\Lambda t} \right)^{\frac{d}{2}}$$

and thus, for all  $v \in \mathbb{R}^d$ ,

$$\int_{|\xi| \leq 1} e^{-\mu_\xi t} |\hat{f}_0|^2 d\xi \leq C \|f_0(\cdot, v)\|_{L^1(dx)}^2 \int_{\mathbb{R}^d} e^{-\frac{\Lambda}{2} |\xi|^2 t} d\xi \leq C \|f_0(\cdot, v)\|_{L^1(dx)}^2 t^{-\frac{d}{2}}.$$

Using the fact that  $\mu_\xi \geq \frac{1}{2}\Lambda$  when  $|\xi| \geq 1$  and Plancherel's formula, we know that, for all  $v \in \mathbb{R}^d$ ,

$$\int_{|\xi| > 1} e^{-\mu_\xi t} |\hat{f}_0|^2 d\xi \leq C e^{-\frac{\Lambda}{2} t} \|f_0(\cdot, v)\|_{L^2(dx)}^2,$$

which completes a first proof of Theorem 1.

## 5. Hypocoercivity and the Nash inequality

In view of the proof of Theorem 1 in Section 4.2 and of the rate, it is natural to wonder if the hypocoercivity can be controlled by the use of Nash's inequality. Here we temporarily abandon the Fourier variable  $\xi$  and consider the direct variable  $x \in \mathbb{R}^d$ : throughout this section, the *transport operator* on the position space is defined as

$$\mathbb{T}f = v \cdot \nabla_x f.$$

We rely on the abstract setting of Section 2, applied to (1) with the scalar product  $\langle \cdot, \cdot \rangle$  on  $L^2(dx d\gamma_\infty)$  and the induced norm  $\|\cdot\|$ . Notice that this norm includes the  $x$  variable, which was not the case in the mode-by-mode analysis of Section 2. It is then easy to check that  $(\mathbb{T}\Pi)f = M\mathbb{T}\rho_f = v \cdot \nabla_x \rho_f M$ ,  $(\mathbb{T}\Pi)^* f = -\nabla_x \cdot (\int_{\mathbb{R}^d} v f dv) M$ , and  $(\mathbb{T}\Pi)^*(\mathbb{T}\Pi)f = -\Theta \Delta_x \rho_f M$  so that

$$g = Af = (1 + (\mathbb{T}\Pi)^* \mathbb{T}\Pi)^{-1} (\mathbb{T}\Pi)^* f \iff g = uM,$$

where  $u - \Theta \Delta u = -\nabla_x \cdot (\int_{\mathbb{R}^d} v f dv)$ . Since  $M$  is unbiased,  $Af = A(1 - \Pi)f$ . For some  $\delta > 0$  to be chosen later, we redefine the entropy by  $H[f] := \frac{1}{2} \|f\|^2 + \delta \langle Af, f \rangle$ .

*Proof of Theorem 1.* If  $f$  solves (1), the time derivative of  $H[f(t, \cdot, \cdot)]$  is given by

$$\frac{d}{dt} H[f] = -D[f], \tag{12}$$

where, as in the proof of Proposition 4,

$$D[f] := -\langle Lf, f \rangle + \delta \langle A\mathbb{T}\Pi f, f \rangle + \delta \operatorname{Re} \langle A\mathbb{T}(1 - \Pi)f, f \rangle - \delta \operatorname{Re} \langle \mathbb{T}A f, f \rangle - \delta \operatorname{Re} \langle A L f, f \rangle.$$

Here we use the fact that  $\langle Af, Lf \rangle = 0$ . The first term in  $D[f]$  satisfies the microscopic coercivity condition

$$-\langle Lf, f \rangle \geq \lambda_m \|(1 - \Pi)f\|^2.$$

The second term in (12) is computed as follows. Solving  $g = A\mathbb{T}\Pi f$  is equivalent to solving  $(1 + (\mathbb{T}\Pi)^*\mathbb{T}\Pi)g = (\mathbb{T}\Pi)^*\mathbb{T}\Pi f$ , i.e.,

$$v_f - \Theta \Delta_x v_f = -\Theta \Delta_x \rho_f, \quad (13)$$

where  $g = v_f M$ . Hence

$$\langle A\mathbb{T}\Pi f, f \rangle = \int_{\mathbb{R}^d} v_f \rho_f dx.$$

A direct application of the hypocoercivity approach of [Dolbeault et al. 2015] to the whole-space problem fails by lack of a *macroscopic coercivity* condition. Although the second term in (12) is not coercive, we observe that the last three terms in (12) can still be dominated by the first two for  $\delta > 0$ , small enough, as follows.

(1) As in [Dolbeault et al. 2015], we use the adjoint operators to compute

$$\langle A\mathbb{T}(1 - \Pi)f, f \rangle = -\langle (1 - \Pi)f, \mathbb{T}A^*f \rangle.$$

We observe that

$$A^*f = \mathbb{T}\Pi(1 + (\mathbb{T}\Pi)^*\mathbb{T}\Pi)^{-1}f = \mathbb{T}(1 + (\mathbb{T}\Pi)^*\mathbb{T}\Pi)^{-1}\Pi f = M\mathbb{T}u_f = vM \cdot \nabla_x u_f,$$

where  $u_f$  is the solution in  $H^1(dx)$  of

$$u_f - \Theta \Delta_x u_f = \rho_f. \quad (14)$$

With  $K$  defined by (11), we obtain

$$\|\mathbb{T}A^*f\|^2 \leq K \|\nabla_x^2 u_f\|_{L^2(dx)}^2 = K \|\Delta_x u_f\|_{L^2(dx)}^2.$$

On the other hand, we observe that  $v_f = -\Theta \Delta_x u_f$  solves (13). Hence by multiplying (14) by  $v_f = -\Theta \Delta_x u_f$  and integrating by parts, we know that

$$\Theta \|\nabla_x u_f\|_{L^2(dx)}^2 + \Theta^2 \|\Delta_x u_f\|_{L^2(dx)}^2 = \int_{\mathbb{R}^d} v_f \rho_f dx = \langle A\mathbb{T}\Pi f, f \rangle. \quad (15)$$

Notice that a central feature of our method is the fact that quantities of interest involving the operator  $A$  can be computed by solving an elliptic equation (for instance (13) in case of  $A\mathbb{T}\Pi f$  or (14) in case of  $A^*f$ ). Altogether we obtain

$$|\langle A\mathbb{T}(1 - \Pi)f, f \rangle| \leq \|(1 - \Pi)f\| \|\mathbb{T}A^*f\| \leq \frac{\sqrt{K}}{\Theta} \|(1 - \Pi)f\| \langle A\mathbb{T}\Pi f, f \rangle^{\frac{1}{2}}.$$

(2) By (8), we have

$$|\langle \mathbb{T}A f, f \rangle| = |\langle \mathbb{T}A(1 - \Pi)f, (1 - \Pi)f \rangle| \leq \|(1 - \Pi)f\|^2.$$

(3) It remains to estimate the last term on the right-hand side of (12). Let us consider the solution  $u_f$  of (14). If we multiply (13) by  $u_f$  and integrate, we observe that

$$\Theta \|\nabla_x u_f\|_{L^2(dx)}^2 = \int_{\mathbb{R}^d} u_f v_f dx \leq \int_{\mathbb{R}^d} u_f v_f dx + \int_{\mathbb{R}^d} |v_f|^2 dx = \int_{\mathbb{R}^d} v_f \rho_f dx$$

because  $v_f = -\Theta \Delta u_f$ , so that

$$\|A^* f\|^2 = \Theta \|\nabla_x u_f\|_{L^2(dx)}^2 \leq \langle \text{AT} \Pi f, f \rangle.$$

In Case (a), we compute

$$\langle \text{AL} f, f \rangle = \langle \text{L}(1 - \Pi) f, A^* f \rangle = \iint_{\mathbb{R}^d \times \mathbb{R}^d} \nabla_x u_f \cdot \frac{\nabla_v M}{M} (1 - \Pi) f dx dv.$$

It follows from the Cauchy–Schwarz inequality that

$$\int_{\mathbb{R}^d} |\nabla_v M| |(1 - \Pi) f| d\gamma_\infty \leq \|\nabla_v M\|_{L^2(d\gamma_\infty)} \|(1 - \Pi) f\|_{L^2(d\gamma_\infty)} = \sqrt{d\theta} \|(1 - \Pi) f\|_{L^2(d\gamma_\infty)}$$

and

$$|\langle \text{AL} f, f \rangle| \leq \|\nabla_x u_f\|_{L^2(dx)} \left( \int_{\mathbb{R}^d} \left( \frac{1}{d} \int_{\mathbb{R}^d} |\nabla_v M| |(1 - \Pi) f| d\gamma \right)^2 dx \right)^{\frac{1}{2}}.$$

Altogether, we obtain

$$|\langle \text{AL} f, f \rangle| \leq \sqrt{\frac{\theta}{\Theta}} \|(1 - \Pi) f\| \langle \text{AT} \Pi f, f \rangle^{\frac{1}{2}}.$$

In Case (b), we use (H2) to get

$$|\langle \text{AL} f, f \rangle| \leq \|L f\| \|A^* f\| \leq 2\bar{\sigma} \|(1 - \Pi) f\| \|A^* f\| \leq 2\bar{\sigma} \|(1 - \Pi) f\| \langle \text{AT} \Pi f, f \rangle^{\frac{1}{2}}.$$

In both cases, (a) and (b), the estimate can be written as

$$|\langle \text{AL} f, f \rangle| \leq 2\bar{\sigma} \|(1 - \Pi) f\| \langle \text{AT} \Pi f, f \rangle^{\frac{1}{2}},$$

with the convention that  $\bar{\sigma} = \frac{1}{2} \sqrt{\theta/\Theta}$  in Case (a).

Summarizing, we know that

$$-\frac{d}{dt} \text{H}[f] \geq (\lambda_m - \delta) X^2 + \delta Y^2 + 2\delta b X Y,$$

with  $X := \|(1 - \Pi) f\|$ ,  $Y := \langle \text{AT} \Pi f, f \rangle^{1/2}$ , and  $b := K/(2\Theta) + 2\bar{\sigma}$ . The largest  $a > 0$  such that

$$(\lambda_m - \delta) X^2 + \delta Y^2 + 2\delta b X Y \geq a(X^2 + 2Y^2)$$

holds for any  $X, Y \in \mathbb{R}$  is given by the conditions

$$a < \lambda_m - \delta, \quad 2a < \delta, \quad \delta^2 b^2 - (\lambda_m - \delta - a)(\delta - 2a) \leq 0 \tag{16}$$

and it is easy to check that there exists a positive solution if  $\delta > 0$  is small enough. To fulfill the additional constraint  $\delta < 1$ , we can for instance choose

$$\delta = \frac{4 \min\{1, \lambda_m\}}{8b^2 + 5} \quad \text{and} \quad a = \frac{\delta}{4}.$$

Altogether we obtain

$$-\frac{d}{dt}H[f] \geq a(\|(1 - \Pi)f\|^2 + 2\langle \text{AT}\Pi f, f \rangle).$$

Using (14) and (15), we control  $\|\Pi f\|^2 = \|\rho_f\|_{L^2(dx)}^2$  by  $\langle \text{AT}\Pi f, f \rangle$  according to

$$\|\Pi f\|^2 = \|u_f\|_{L^2(dx)}^2 + 2\Theta\|\nabla_x u_f\|_{L^2(dx)}^2 + \Theta^2\|\Delta_x u_f\|_{L^2(dx)}^2 \leq \|u_f\|_{L^2(dx)}^2 + 2\langle \text{AT}\Pi f, f \rangle.$$

We observe that, for any  $t \geq 0$ ,

$$\|u_f(t, \cdot)\|_{L^1(dx)} = \|\rho_f(t, \cdot)\|_{L^1(dx)} = \|f_0\|_{L^1(dx dv)}, \quad \|\nabla_x u_f\|_{L^2(dx)}^2 \leq \frac{1}{\Theta} \langle \text{AT}\Pi f, f \rangle.$$

We recall the *Nash inequality* [1958]

$$\|u\|_{L^2(dx)}^2 \leq \mathcal{C}_{\text{Nash}} \|u\|_{L^1(dx)}^{\frac{4}{d+2}} \|\nabla u\|_{L^2(dx)}^{\frac{2d}{d+2}} \quad (17)$$

for any function  $u \in L^1 \cap H^1(\mathbb{R}^d)$ . We use (17) with  $u = u_f$  to get

$$\|\Pi f\|^2 \leq \Phi^{-1}(2\langle \text{AT}\Pi f, f \rangle), \quad \text{with} \quad \Phi^{-1}(y) := y + \left(\frac{y}{c}\right)^{\frac{d}{d+2}} \quad \text{for all } y \geq 0,$$

where

$$c = 2\Theta \mathcal{C}_{\text{Nash}}^{-1-\frac{2}{d}} \|f_0\|_{L^1(dx dv)}^{-\frac{4}{d}}.$$

The function  $\Phi : [0, \infty) \rightarrow [0, \infty)$  satisfies  $\Phi(0) = 0$  and  $0 < \Phi' < 1$ , so that

$$\|(1 - \Pi)f\|^2 + 2\langle \text{AT}\Pi f, f \rangle \geq \Phi(\|f\|^2) \geq \Phi\left(\frac{2}{1+\delta}H[f]\right),$$

where the last inequality holds as a consequence of (9). From

$$z = \Phi^{-1}(y) = y + \left(\frac{y}{c}\right)^{\frac{d}{d+2}} \leq y_0^{\frac{2}{d+2}} y^{\frac{d}{d+2}} + \left(\frac{y}{c}\right)^{\frac{d}{d+2}} = (y_0^{\frac{2}{d+2}} + c^{-\frac{d}{d+2}})y^{\frac{d}{d+2}},$$

as long as  $y \leq y_0$ , for  $y_0$  to be chosen later, we have

$$y = \Phi(z) \geq (\Phi(z_0)^{\frac{2}{d+2}} + c^{-\frac{d}{d+2}})^{-\frac{d+2}{d}} z^{1+\frac{2}{d}},$$

as long as  $z \leq z_0 := \Phi^{-1}(y_0)$ . Since  $\frac{d}{dt}H[f] \leq 0$ , we have

$$\frac{2}{1+\delta}H[f] \leq \frac{2}{1+\delta}H[f_0].$$

We thus apply the previous inequalities with

$$z_0 = \frac{2}{1+\delta}H[f_0]$$

together with the fact that

$$\Phi(z_0) \geq z_0 \geq \frac{1-\delta}{1+\delta} \|f_0\|^2$$

and that  $c$  is proportional to  $\|f_0\|_{L^1(dx dv)}^{-4/d}$ , to get

$$\Phi\left(\frac{2}{1+\delta} H[f]\right) \gtrsim (\|f_0\|_{L^2(dx d\gamma_\infty)}^{\frac{4}{d+2}} + \|f_0\|_{L^1(dx dv)}^{\frac{4}{d+2}})^{-\frac{d+2}{d}} H[f]^{1+\frac{2}{d}}.$$

We deduce the entropy decay inequality

$$-\frac{d}{dt} H[f] \gtrsim (\|f_0\|_{L^2(dx d\gamma_\infty)}^{\frac{4}{d+2}} + \|f_0\|_{L^1(dx dv)}^{\frac{4}{d+2}})^{-\frac{d+2}{d}} H[f]^{1+\frac{2}{d}}. \quad (18)$$

A simple integration from 0 to  $t$  shows that

$$H[f] \lesssim [H[f_0]^{-\frac{2}{d}} + (\|f_0\|_{L^2(dx d\gamma_\infty)}^{\frac{4}{d+2}} + \|f_0\|_{L^1(dx dv)}^{\frac{4}{d+2}})^{-\frac{d+2}{d}} t]^{-\frac{d}{2}}.$$

The result of Theorem 1 then follows from elementary considerations.  $\square$

Using moments instead of the mass, it is possible to state an *improved Nash inequality*: there exists a positive constant  $\mathcal{C}_\star$  such that

$$\|u\|_{L^2(dx)}^2 \leq \mathcal{C}_\star \|xu\|_{L^1(dx)}^{\frac{4}{d+4}} \|\nabla u\|_{L^2(dx)}^{\frac{d+2}{d+4}}$$

for any  $u \in H^1(dx) \cap L^1((1+|x|)dx)$  such that  $\int_{\mathbb{R}^d} u dx = 0$ . The proof follows from a minor modification of the original proof (attributed by Nash himself to Stein) in [Nash 1958] and uses Fourier variables. As a consequence, any solution of the heat equation with zero average decays in  $L^2(dx)$  like  $O(t^{-1-d/2})$  as  $t \rightarrow +\infty$ . It is the topic of the following section to use Fourier variables in the spirit of Nash's proof to get improved rates of decay at the level of the kinetic equation.

## 6. Algebraic decay rates in $\mathbb{R}^d$ by Fourier estimates and improvements

We prove Theorem 2 in Section 6.1 and Theorem 3 in Section 6.2.

**6.1. Improved decay rates.** Let us prove Theorem 2 by Fourier methods inspired by the proof of Nash's inequality.

Step 1: decay of the average in space by a factorization argument. We define

$$f_\bullet(t, v) := \int_{\mathbb{R}^d} f(t, x, v) dx \quad (19)$$

and observe that  $f_\bullet$  solves

$$\partial_t f_\bullet = L f_\bullet.$$

As a consequence, we have that  $0 = \int_{\mathbb{R}^d} f_\bullet(t, v) dv$ . From the *microscopic coercivity property* (H4), we deduce that

$$\|f_\bullet(t, \cdot)\|_{L^2(d\gamma_\infty)}^2 = \int_{\mathbb{R}^d} \left| \frac{f_\bullet(t, v)}{M} \right|^2 M dv \leq \|f_\bullet(0, \cdot)\|_{L^2(d\gamma_\infty)}^2 e^{-\lambda_m t} \quad \text{for all } t \geq 0.$$

With  $k \in (d, \infty)$ , Proposition 6 applies like in the proof of Corollary 8 or in [Mischler and Mouhot 2016]. We observe that  $\|f_\bullet(0, \cdot)\|_{L^2(|v|^2 d\gamma_k)} \leq \|f_0\|_{L^2(|v|^2 d\gamma_k; L^1(dx))}$ . For some positive constants  $C$  and  $\lambda$ , we get

$$\|f_\bullet(t, \cdot)\|_{L^2(|v|^2 d\gamma_k)}^2 \leq C \|f_0\|_{L^2(|v|^2 d\gamma_k; L^1(dx))}^2 e^{-\lambda t} \quad \text{for all } t \geq 0. \quad (20)$$

**Step 2:** improved decay of  $f$ . Let us define  $g(t, x, v) := f(t, x, v) - f_\bullet(t, v)\varphi(x)$ , where  $\varphi$  is a given positive function satisfying

$$\int_{\mathbb{R}^d} \varphi(x) dx = 1, \quad \text{e.g., } \varphi(x) := (2\pi)^{-\frac{d}{2}} e^{-\frac{|x|^2}{2}}, \quad \text{for all } x \in \mathbb{R}^d.$$

Since  $\partial_t f_\bullet = Lf_\bullet$ , the Fourier transform  $\hat{g}(t, \xi, v)$  of  $g(t, x, v)$  solves

$$\partial_t \hat{g} + \mathbb{T} \hat{g} = L \hat{g} - f_\bullet \mathbb{T} \hat{\varphi},$$

where  $\mathbb{T} \hat{\varphi} = i(v \cdot \xi) \hat{\varphi}$ . Using Duhamel's formula

$$\hat{g} = e^{(L-\mathbb{T})t} \hat{g}_0 - \int_0^t e^{(L-\mathbb{T})(t-s)} f_\bullet(s, v) \mathbb{T} \hat{\varphi}(\xi) ds,$$

Corollary 5, and Proposition 6, for some generic constant  $C > 0$  which will change from line to line, we get

$$\|\hat{g}(t, \xi, \cdot)\|_{L^2(d\gamma_k)} \leq C e^{-\frac{1}{2}\mu_\xi t} \|\hat{g}_0(\xi, \cdot)\|_{L^2(d\gamma_k)} + C \int_0^t e^{-\frac{\mu_\xi}{2}(t-s)} \|f_\bullet(s, \cdot)\|_{L^2(|v|^2 d\gamma_k)} |\xi| |\hat{\varphi}(\xi)| ds. \quad (21)$$

The key observation is  $\hat{g}_0(0, v) = 0$ , so that

$$\hat{g}_0(\xi, v) = \int_0^{|\xi|} \frac{\xi}{|\xi|} \cdot \nabla_\eta \hat{g}_0\left(\eta \frac{\xi}{|\xi|}, v\right) d\eta$$

yields

$$|\hat{g}_0(\xi, v)| \leq |\xi| \|\nabla_\xi \hat{g}_0(\cdot, v)\|_{L^\infty(d\xi)} \leq |\xi| \|g_0(\cdot, v)\|_{L^1(|x| dx)} \quad \text{for all } (\xi, v) \in \mathbb{R}^d \times \mathbb{R}^d.$$

We know from (10) that  $\mu_\xi = \Lambda |\xi|^2 / (1 + |\xi|^2)$ . The first term of the right-hand side of (21) can therefore be estimated for any  $t \geq 1$  by

$$\begin{aligned} \left( \int_{|\xi| \leq 1} \int_{\mathbb{R}^d} |e^{(L-\mathbb{T})t} \hat{g}_0|^2 d\gamma_k d\xi \right)^{\frac{1}{2}} &\leq \left( \int_{\mathbb{R}^d} |\xi|^2 e^{-\frac{\Lambda}{2} |\xi|^2 t} d\xi \right)^{\frac{1}{2}} \|g_0\|_{L^2(d\gamma_k; L^1(|x| dx))} \\ &\leq \frac{C}{(1+t)^{1+\frac{d}{2}}} \|g_0\|_{L^2(d\gamma_k; L^1(|x| dx))}, \end{aligned}$$

which is the leading-order term as  $t \rightarrow \infty$ , and we have

$$\int_{|\xi| > 1} e^{-\mu_\xi t} \|\hat{g}_0(\xi, \cdot)\|_{L^2(d\gamma_k)}^2 d\xi \leq C e^{-\frac{\Lambda}{2} t} \|g_0\|_{L^2(dx d\gamma_k)}^2$$

for any  $t \geq 0$ , using the fact that  $\mu_\xi \geq \frac{1}{2}\Lambda$  when  $|\xi| \geq 1$  and Plancherel's formula.

Using (20), the second term of the right-hand side of (21) is estimated by

$$\begin{aligned} & \int_{\mathbb{R}^d} \left( \int_0^t e^{-\frac{\mu_\xi}{2}(t-s)} \|f_\bullet(s, \cdot)\|_{L^2(|v|^2 d\gamma_k)} |\xi| |\hat{\varphi}(\xi)| ds \right)^2 d\xi \\ & \leq C \|f_0\|_{L^2(|v|^2 d\gamma_k; L^1(dx))}^2 \int_{\mathbb{R}^d} |\xi|^2 |\hat{\varphi}(\xi)|^2 \left( \int_0^t e^{-\frac{\mu_\xi}{2}(t-s)} e^{-\frac{\lambda}{2}s} ds \right)^2 d\xi. \end{aligned}$$

On the one hand, we use the Cauchy–Schwarz inequality to get

$$\begin{aligned} & \int_{|\xi| \leq 1} |\xi|^2 |\hat{\varphi}(\xi)|^2 \left( \int_0^t e^{-\frac{\mu_\xi}{2}(t-s)} e^{-\frac{\lambda}{2}s} ds \right)^2 d\xi \\ & \leq \|\varphi\|_{L^1(dx)}^2 \int_{|\xi| \leq 1} |\xi|^2 \left( \int_0^t e^{-\mu_\xi(t-s)} e^{-\frac{\lambda}{2}s} ds \right) \left( \int_0^t e^{-\frac{\lambda}{2}s} ds \right) d\xi \\ & \leq \frac{2}{\lambda} \|\varphi\|_{L^1(dx)}^2 \int_0^t \left( \int_{|\xi| \leq 1} |\xi|^2 e^{-\frac{\Lambda}{2}|\xi|^2(t-s)} d\xi \right) e^{-\frac{\lambda}{2}s} ds \leq C_1 t^{-\frac{d}{2}-1} + C_2 e^{-\frac{\lambda}{4}t}, \end{aligned}$$

where the last inequality is obtained by splitting the integral in  $s$  on  $(0, \frac{1}{2}t)$  and  $(\frac{1}{2}t, t)$ . On the other hand, using  $\mu_\xi \geq \frac{1}{2}\Lambda$  when  $|\xi| \geq 1$ , we obtain

$$\int_{|\xi| \geq 1} |\xi|^2 |\hat{\varphi}(\xi)|^2 \left( \int_0^t e^{-\frac{\mu_\xi}{2}(t-s)} e^{-\frac{\lambda}{2}s} ds \right)^2 d\xi \leq t^2 e^{-\min\{\frac{1}{2}\Lambda, \lambda\}t} \|\nabla\varphi\|_{L^2(dx)}^2.$$

By collecting all terms, we deduce that  $\|g(t, \cdot, \cdot)\|_{L^2(dx d\gamma_k)}^2$  is bounded by

$$C(\|g_0\|_{L^2(d\gamma_k; L^1(|x| dx))}^2 + \|f_0\|_{L^2(|v|^2 d\gamma_k; L^1(dx))}^2)(1+t)^{-(1+\frac{d}{2})}$$

for some constant  $C > 0$ . Recalling that  $f = g + f_\bullet\varphi$ , the proof of Theorem 2 is completed using (20).

**6.2. Improved decay rates with higher-order cancellations.** We prove Theorem 3, which means that from now on we assume in Case (a) that  $M$  is a normalized Gaussian (2), and in Case (b) that  $\sigma \equiv 1$ . Moreover, the initial data satisfies (5); that is,

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} f_0 P dx dv = 0 \quad \text{for all } P \in \mathbb{R}_\ell[X, V].$$

For any  $P \in \mathbb{R}_\ell[X]$ , let

$$P[f](t, v) := \int_{\mathbb{R}^d} P(x) f(t, x, v) dx,$$

so that  $\int_{\mathbb{R}^d} P[f](0, v) dv = 0$ .

In this section we use the notation  $\lesssim_k$  to express inequalities up to a constant which depends on  $k$ .

**Step 1:** conservation of zero moments. For a solution  $f$  of (1) we compute

$$\begin{aligned} \frac{d}{dt} \iint_{\mathbb{R}^d \times \mathbb{R}^d} f(t, x, v) P(x, v) dx dv &= - \iint_{\mathbb{R}^d \times \mathbb{R}^d} (v \cdot \nabla_x f) P dx dv + \iint_{\mathbb{R}^d \times \mathbb{R}^d} (\mathbb{L}f) P dx dv \\ &= \iint_{\mathbb{R}^d \times \mathbb{R}^d} (v \cdot \nabla_x P) f dx dv + \iint_{\mathbb{R}^d \times \mathbb{R}^d} (\mathbb{L}f) P dx dv. \end{aligned}$$

In Case (a) of a Fokker–Planck operator, we may write

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} (\mathsf{L}f)P \, dx \, dv = \iint_{\mathbb{R}^d \times \mathbb{R}^d} \frac{1}{M} \nabla_v \cdot (M \nabla_v P) f \, dx \, dv = \iint_{\mathbb{R}^d \times \mathbb{R}^d} (\Delta_v P - v \cdot \nabla_v P) f \, dx \, dv.$$

By the definition of  $\mathbb{R}_\ell[X, V]$ , it turns out that  $\Delta_v P - v \cdot \nabla_v P \in \mathbb{R}_\ell[X, V]$ . For the scattering operator of Case (b), one has

$$\begin{aligned} \iint_{\mathbb{R}^d \times \mathbb{R}^d} (\mathsf{L}f)P \, dx \, dv &= \iint_{\mathbb{R}^d \times \mathbb{R}^d} \left( \int_{\mathbb{R}^d} (M(v) f(t, x, v') - M(v') f(t, x, v)) \, dv' \right) P(x, v) \, dx \, dv \\ &= \iiint_{\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d} (M(v) f(t, x, v') - M(v') f(t, x, v)) P(x, v) \, dx \, dv \, dv' \\ &= \iint_{\mathbb{R}^d \times \mathbb{R}^d} \left( \int_{\mathbb{R}^d} M(v) P(x, v) \, dv \right) f(t, x, v') \, dx \, dv' - \iint_{\mathbb{R}^d \times \mathbb{R}^d} f(x, v) P(x, v) \, dx \, dv. \end{aligned}$$

One can check that  $\int_{\mathbb{R}^d} M(v) P(x, v) \, dv \in \mathbb{R}_\ell[X]$ . Since also  $v \cdot \nabla_x P \in \mathbb{R}_\ell[X, V]$ , the evolution of moments of order lower than or equal to  $\ell$  is equivalent to a linear ODE of the form  $\dot{Y}(t) = QY(t)$ , where  $Q$  is a matrix resulting from the previous computations. Consequently, if  $Y(0) = 0$  initially, it remains null for all times.

**Step 2:** decay of polynomial averages in space. We claim that for any  $j \leq \ell$ , there exists  $\lambda > 0$  such that, for any  $P \in \mathbb{R}_j[X]$  and  $q \in \mathbb{N}$ ,

$$\|P[f](t, \cdot)\|_{L^2(d\gamma_{k+q})} \lesssim_{j,q} \|f_0\|_{L^2(d\gamma_{k+q+2j}; L^1((1+|x|^j) dx))} (1+t)^j e^{-\lambda t} \quad \text{for all } t \geq 0. \quad (22)$$

Let us prove it by induction.

(1) *The case  $j = 0$ .* Notice that  $j = 0$  means that  $P$  is a real number and  $P[f] = f$ , as defined in (19), up to a multiplication by a constant. Since  $\int_{\mathbb{R}^d} f_\bullet(t, v) \, dv = 0$  for any  $t \geq 0$ , one has  $\partial_t f_\bullet = \mathsf{L}f_\bullet$ ; thus we deduce from the *microscopic coercivity property* as above that

$$\|f_\bullet(t, \cdot)\|_{L^2(d\gamma_\infty)} \leq \|f_\bullet(0, \cdot)\|_{L^2(d\gamma_\infty)} e^{-\lambda_m t} \quad \text{for all } t \geq 0.$$

We also obtain

$$\|f_\bullet(t, \cdot)\|_{L^2(d\gamma_{k+q})} \lesssim_q \|f_0\|_{L^2(d\gamma_{k+q}; L^1(dx))} e^{-\lambda t} \quad \text{for all } t \geq 0, \quad (23)$$

but this requires some comments. The case  $k \in (d, \infty)$  is covered by Corollary 8.

The case  $k = \infty$  in (23) is given by the following lemma.

**Lemma 10.** *Under the assumptions of Theorem 3, one has*

$$\|f_\bullet(t, \cdot)\|_{L^2((1+|v|^q) d\gamma_\infty)} \lesssim_q \|f_0\|_{L^2((1+|v|^q) d\gamma_\infty; L^1(dx))} e^{-\lambda t} \quad \text{for all } t \geq 0.$$

*Proof.* We rely on Proposition 7 with the Banach spaces  $\mathcal{B}_1 = L^2(d\gamma_\infty)$  and  $\mathcal{B}_2 = L^2((1+|v|^q) d\gamma_\infty)$ . In Case (a), let us define  $\mathfrak{A}$  and  $\mathfrak{B}$  by  $\mathfrak{A}F = N\chi_R F$  and  $\mathfrak{B}F = \mathsf{L}F - \mathfrak{A}F$ . In Case (b), we consider  $\mathfrak{A}$

and  $\mathfrak{B}$  such that

$$\begin{aligned}\mathfrak{A}F(v) &= M(v) \int_{\mathbb{R}^d} F(v') dv', \\ \mathfrak{B}F(v) &= - \int_{\mathbb{R}^d} M(v') dv' F(v).\end{aligned}$$

The semigroup generated by  $\mathfrak{A} + \mathfrak{B}$  is exponentially decreasing in  $\mathcal{B}_1$  by the microscopic coercivity property, as above. The semigroup generated by  $\mathfrak{B}$  is exponentially decreasing in  $\mathcal{B}_2$ . In Case (b), it is straightforward. In Case (a),  $F(t) = e^{\mathfrak{B}t} F_0$  is such that

$$\begin{aligned}& \frac{1}{2} \frac{d}{dt} \int_{\mathbb{R}^d} |F|^2 (1+|v|^q) d\gamma_\infty \\ &= \int_{\mathbb{R}^d} (\mathfrak{B}F) F (1+|v|^q) d\gamma_\infty \\ &= \int_{\mathbb{R}^d} \nabla_v \cdot \left( M \nabla_v \left( \frac{F}{M} \right) \right) F (1+|v|^q) d\gamma_\infty - \int_{\mathbb{R}^d} N \chi_R(v) |F|^2 (1+|v|^q) d\gamma_\infty \\ &= - \int_{\mathbb{R}^d} \left| \nabla_v \left( \frac{F}{M} \right) \right|^2 (1+|v|^q) M dv - \int_{\mathbb{R}^d} q |v|^{q-2} v \cdot \nabla_v \left( \frac{F}{M} \right) \frac{F}{M} M dv - \int_{\mathbb{R}^d} N \chi_R(v) |F|^2 (1+|v|^q) \frac{dv}{M} \\ &\leq \int_{\mathbb{R}^d} \left\{ \frac{q}{2} \frac{\nabla_v \cdot (|v|^{q-2} v M)}{(1+|v|^q) M} - N \chi_R(v) \right\} |F|^2 (1+|v|^q) \frac{dv}{M} \leq -\frac{\lambda}{2} \int_{\mathbb{R}^d} |F|^2 (1+|v|^q) d\gamma_\infty\end{aligned}$$

for some  $\lambda > 0$ , by choosing  $N$  and  $R$  large enough.

The operator  $\mathfrak{A} : \mathcal{B}_1 \rightarrow \mathcal{B}_2$  is bounded. This is straightforward in Case (a) and follows from the boundedness of  $\int_{\mathbb{R}^d} M(v)(1+|v|^q) d\gamma_\infty$  in Case (b). Proposition 7 applies, which concludes the proof.  $\square$

(2) *Induction.* Let us assume that (22) is true for some  $j \geq 0$ , consider  $P \in \mathbb{R}_{j+1}[X]$ , and observe that  $P[f]$  solves

$$\partial_t P[f] = \mathbb{L}P[f] - \int_{\mathbb{R}^d} (v \cdot \nabla_x P) f dx.$$

Since  $\nabla_x P \in \mathbb{R}_j[X]$ , the induction hypothesis at step  $j$  (applied with  $q$  replaced by  $q + 2$ ) gives

$$\begin{aligned}\left\| v \int_{\mathbb{R}^d} (\nabla_x P)[f] dx \right\|_{L^2(d\gamma_{k+q})} &\lesssim \left\| \int_{\mathbb{R}^d} (\nabla_x P)[f] dx \right\|_{L^2(d\gamma_{k+q+2})} \\ &\lesssim_{j,q} \|f_0\|_{L^2(d\gamma_{k+q+2(j+1)}; L^1((1+|x|^j) dx))} (1+t)^j e^{-\lambda t}.\end{aligned}$$

By Duhamel's formula, we have

$$P[f](t, v) = e^{\mathbb{L}t} P[f](0, v) - \int_0^t e^{\mathbb{L}(t-s)} \left( v \int_{\mathbb{R}^d} (\nabla_x P)[f_s] dx \right) ds.$$

Note that

$$\int_{\mathbb{R}^d} v \int_{\mathbb{R}^d} (\nabla_x P)[f] dx dv = \iint_{\mathbb{R}^d \times \mathbb{R}^d} (v \cdot \nabla_x P)[f] dx dv = 0$$

for all  $t \geq 0$  since  $v \cdot \nabla_x P \in \mathbb{R}_\ell[X, V]$ . As a consequence, the decay of the semigroup associated with  $L$  can be estimated by

$$\left\| e^{L(t-s)} \left( v \int_{\mathbb{R}^d} (\nabla_x P)[f_s] dx \right) \right\|_{L^2(d\gamma_\infty)} \leq \left\| v \int_{\mathbb{R}^d} (\nabla_x P)[f_s] dx \right\|_{L^2(d\gamma_\infty)} e^{-\lambda_m(t-s)}.$$

As in the case  $j = 0$ , we deduce from Corollary 8 that

$$\begin{aligned} \left\| e^{L(t-s)} \left( v \int_{\mathbb{R}^d} (\nabla_x P)[f_s] dx \right) \right\|_{L^2((1+|v|^q) d\gamma_k)} &\leq \left\| v \int_{\mathbb{R}^d} (\nabla_x P)[f_s] dx \right\|_{L^2(d\gamma_{k+q})} e^{-\lambda(t-s)} \\ &\lesssim_{q,k} \|f_0\|_{L^2(d\gamma_{k+q+2(j+1)}; L^1((1+|x|^j) dx))} (1+s)^j e^{-\lambda t}. \end{aligned}$$

Moreover, since

$$\iint_{\mathbb{R}^d \times \mathbb{R}^d} f_0(x, v) P(x) dx dv = 0,$$

for the same reasons we also have

$$\|e^{Lt} P[f](0, \cdot)\|_{L^2(d\gamma_{k+q})} \leq \|P[f_0]\|_{L^2((1+|v|^q) d\gamma_k)} e^{-\lambda t}$$

for some  $\lambda > 0$ . We deduce from Duhamel's formula that

$$\begin{aligned} \|P[f]\|_{L^2(d\gamma_{k+q})} &\lesssim \|e^{Lt} P[f](0, \cdot)\|_{L^2(d\gamma_{k+q})} + \int_0^t \left\| e^{-L(t-s)} \left( v \int_{\mathbb{R}^d} \nabla_x P[f_s] dx \right) \right\|_{L^2(d\gamma_{k+q})} ds \\ &\lesssim_k \|f_0\|_{L^2(d\gamma_{k+q}; L^1((1+|x|^{j+1}) dx))} e^{-\lambda t} + \int_0^t (1+s)^j e^{-\lambda t} \|f_0\|_{L^2(d\gamma_{k+q+2(j+1)}; L^1((1+|x|^j) dx))} ds \\ &\lesssim_k \|f_0\|_{L^2(d\gamma_{k+q+2(j+1)}; L^1((1+|x|^{j+1}) dx))} (1+t)^{j+1} e^{-\lambda t}, \end{aligned}$$

which proves the induction.

**Step 3:** improved decay of  $f$ . Let us choose some  $t_0 > 0$ . In order to estimate

$$\|f(t, \cdot, \cdot)\|_{L^2(dx d\gamma_k)}^2 = \|e^{(L-\mathbb{T})t} f_0\|_{L^2(dx d\gamma_k)}^2,$$

we compute its evolution on  $(0, 2t_0)$  and split the interval on  $(0, t_0)$  and  $(t_0, 2t_0)$  using the semigroup property

$$\|e^{(L-\mathbb{T})(2t_0)} f_0\|_{L^2(dx d\gamma_k)}^2 = \|e^{(L-\mathbb{T})t_0} (e^{(L-\mathbb{T})t_0} f_0)\|_{L^2(dx d\gamma_k)}^2.$$

Up to the end of this section,  $\mathbb{T} = v \cdot \nabla_x$  denotes the transport operator in position and velocity variables. We decompose  $f_{t_0} = e^{(L-\mathbb{T})t_0} f_0$  into

$$f_{t_0} = \left( \sum_{|\alpha| \leq \ell} \frac{1}{\alpha!} X^\alpha [f_{t_0}] \partial^\alpha \varphi \right) + g_0, \quad \text{with } g_0 := f_{t_0} - \sum_{|\alpha| \leq \ell} \frac{1}{\alpha!} X^\alpha [f_{t_0}] \partial^\alpha \varphi,$$

where  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_i, \dots, \alpha_d) \in \mathbb{N}^d$  is a multi-index such that  $|\alpha| = \sum_{i=1}^d \alpha_i \leq \ell$  and  $\varphi$  is given by

$$\varphi(x) := (2\pi)^{-\frac{d}{2}} e^{-\frac{|x|^2}{2}} \quad \text{for all } x \in \mathbb{R}^d.$$

Here we use the notation  $\partial^\alpha \varphi = \partial_{x_1}^{\alpha_1} \partial_{x_2}^{\alpha_2} \cdots \partial_{x_d}^{\alpha_d} \varphi$  and  $X^\alpha = \prod_i^n X_i^{\alpha_i}$ . According to (22), we know that

$$\|X^\alpha [f_{t_0}]\|_{L^2(d\gamma_k)} \lesssim_j \|f_0\|_{L^2(d\gamma_{k+2j}; L^1((1+|x|^j) dx))} (1+t_0)^j e^{-\lambda t_0},$$

so that, by considering the evolution of the first term on  $(t_0, 2t_0)$ , we obtain

$$\left\| e^{(L-\mathbb{T})t_0} \left( \sum_{|\alpha| \leq \ell} \frac{1}{\alpha!} X^\alpha [f_{t_0}] \partial^\alpha \varphi \right) \right\|_{L^2(dx d\gamma_k)} \lesssim \sum_{|\alpha| \leq \ell} \|X^\alpha [f_{t_0}]\|_{L^2(d\gamma_k)} \|\partial^\alpha \varphi\|_{L^2(dx)} \lesssim e^{-\frac{\lambda}{2} t_0}. \quad (24)$$

Next, let us consider the second term and define, on  $t+t_0 \in (t_0, 2t_0)$ , the function

$$g := f_{t+t_0} - \sum_{|\alpha| \leq \ell} \frac{1}{\alpha!} X^\alpha [f_{t+t_0}] \partial^\alpha \varphi.$$

With initial datum  $g_0$ , it solves on  $(0, t_0)$  the equation

$$\begin{aligned} \partial_t g &= \partial_t f_{t+t_0} - \sum_{|\alpha| \leq \ell} \frac{1}{\alpha!} \partial_t (X^\alpha [f_{t+t_0}]) \partial^\alpha \varphi \\ &= (L-\mathbb{T})(f_{t+t_0}) - L \left( \sum_{|\alpha| \leq \ell} \frac{1}{\alpha!} X^\alpha [f_{t+t_0}] \partial^\alpha \varphi \right) + \sum_{|\alpha| \leq \ell} \frac{1}{\alpha!} \left( \int_{\mathbb{R}^d} (v \cdot \nabla_x x^\alpha) f_{t+t_0} dx \right) \partial^\alpha \varphi \\ &= (L-\mathbb{T})(g) - \mathbb{T} \left( \sum_{|\alpha| \leq \ell} \frac{1}{\alpha!} X^\alpha [f_{t+t_0}] \partial^\alpha \varphi \right) + \sum_{|\alpha| \leq \ell} \frac{1}{\alpha!} \left( \int_{\mathbb{R}^d} (v \cdot \nabla_x x^\alpha) f_{t+t_0} dx \right) \partial^\alpha \varphi \\ &= (L-\mathbb{T})(g) + v \sum_{|\alpha| \leq \ell} \frac{1}{\alpha!} (\nabla_x X^\alpha [f] \partial^\alpha \varphi - X^\alpha [f_{t+t_0}] \nabla_x (\partial^\alpha \varphi)), \end{aligned}$$

where  $\alpha! = \prod_{i=1}^d \alpha_i!$  is associated with the multi-index  $\alpha = (\alpha_i)_{i=1}^d$  and

$$\nabla_x X^\alpha [f] = (\partial_{x_i} X^\alpha [f])_{i=1}^d := \left( \int_{\mathbb{R}^d} \partial_{x_i} x^\alpha f dx \right)_{i=1}^d = \left( \int_{\mathbb{R}^d} \alpha_i x^{\alpha \wedge i} f dx \right)_{i=1}^d.$$

Here the notation  $\alpha \wedge i$  denotes the multi-index  $(\alpha_1, \alpha_2, \dots, \alpha_{i-1}, \alpha_i - 1, \alpha_{i+1}, \dots, \alpha_d)$  with the convention that  $X^{\alpha \wedge i} \equiv 0$  if  $\alpha_i = 0$ . We also define the opposite transformation

$$\alpha \vee i := (\alpha_1, \alpha_2, \dots, \alpha_{i-1}, \alpha_i + 1, \alpha_{i+1}, \dots, \alpha_d)$$

so that  $\partial_{x_i} (\partial^\alpha \varphi) = \partial^{\alpha \vee i} \varphi$ . Let us consider the last term and start with the case  $d = 1$ . In that case,

$$\begin{aligned} v \sum_{|\alpha| \leq \ell} \frac{1}{\alpha!} (\nabla_x X^\alpha [f] \partial^\alpha \varphi - X^\alpha [f_{t+t_0}] \nabla_x (\partial^\alpha \varphi)) \\ &= v_1 \sum_{\alpha_1=0}^{\ell} \frac{1}{\alpha_1!} \left( \left( \int_{\mathbb{R}} (\alpha_1 x^{\alpha_1-1}) f_{t+t_0} dx \right) \partial x_1^{\alpha_1} \varphi - \left( \int_{\mathbb{R}} x^{\alpha_1} f_{t+t_0} dx \right) \partial x_1^{\alpha_1+1} \varphi \right) \\ &= -\frac{v_1}{\ell!} \left( \int_{\mathbb{R}} x^\ell f_{t+t_0} dx \right) \partial x_1^{\ell+1} \varphi \end{aligned}$$

because it is a telescoping sum. We adopt the convention that  $\alpha! = 1$  if  $\alpha_i \leq 0$  for some  $i = 1, 2, \dots, d$ . The same property holds in higher dimensions:

$$\begin{aligned} & \sum_{|\alpha| \leq \ell} \frac{1}{\alpha!} (\partial_{x_i} X^\alpha [f] \partial^\alpha \varphi - X^\alpha [f_{t+t_0}] \partial_{x_i} (\partial^\alpha \varphi)) \\ &= \sum_{|\alpha| \leq \ell} \left( \frac{1}{\alpha_{\wedge i}!} X^{\alpha_{\wedge i}} [f] \partial^\alpha \varphi - \frac{1}{\alpha!} X^\alpha [f_{t+t_0}] \partial^{\alpha_{\vee i}} \varphi \right) = - \sum_{|\alpha| = \ell} \frac{1}{\alpha!} X^\alpha [f_{t+t_0}] \partial_{x_i} (\partial^\alpha \varphi). \end{aligned}$$

We deduce that

$$\partial_t g = (L - \mathbb{T})(g) - v \sum_{|\alpha| = \ell} \frac{1}{\alpha!} X^\alpha [f_{t+t_0}] \nabla_x (\partial^\alpha \varphi).$$

Duhamel's formula in Fourier variables gives

$$\hat{g}(t_0, \xi, v) = e^{(L-\mathbb{T})t_0} \hat{g}_0 - \int_0^{t_0} e^{(L-\mathbb{T})(t_0-s)} \left( v \sum_{|\alpha| = \ell} \frac{1}{\alpha!} X^\alpha [f_{s+t_0}] \widehat{\nabla_x (\partial^\alpha \varphi)} \right) ds$$

up to a straightforward abuse of notation. Hence

$$\begin{aligned} & \|\hat{g}(t_0, \xi, \cdot)\|_{L^2(d\gamma_k)} \\ & \lesssim e^{-\frac{1}{2}\mu_\xi t_0} \|\hat{g}_0(\xi, \cdot)\|_{L^2(d\gamma_k)} + \int_0^{t_0} e^{-\frac{\mu_\xi}{2}(t_0-s)} \sum_{|\alpha| = \ell} \frac{1}{\alpha!} \|X^\alpha [f_{s+t_0}]\|_{L^2(|v|^2 d\gamma_k)} |\widehat{\nabla_x (\partial^\alpha \varphi)}| ds. \end{aligned}$$

Recall that (22) gives

$$\|X^\alpha [f_{s+t_0}]\|_{L^2(|v|^2 d\gamma_k)} \lesssim \ell \|f_0\|_{L^2(d\gamma_{k+2\ell+2}; L^1((1+|x|^\ell) dx))} e^{-\frac{\lambda}{2}s}.$$

On the other hand we use  $|\widehat{\nabla_x (\partial^\alpha \varphi)}| \leq |\xi|^{\ell+1} |\hat{\varphi}|$  and observe that

$$|\hat{g}_0(\xi, v)| \lesssim |\xi|^{\ell+1} \|g_0(\cdot, v)\|_{L^1(|x|^\ell dx)} \quad \text{for all } (\xi, v) \in \mathbb{R}^d \times \mathbb{R}^d.$$

Collecting terms, we have

$$\begin{aligned} \|\hat{g}(t_0, \xi, \cdot)\|_{L^2(d\gamma_k)} & \lesssim e^{-\frac{1}{2}\mu_\xi t_0} |\xi|^{\ell+1} \mathbf{1}_{|\xi| < 1} \|g_0(\cdot, v)\|_{L^2(d\gamma_k; L^1(|x|^\ell dx))} + e^{-\frac{1}{2}\mu_\xi t_0} \mathbf{1}_{|\xi| \geq 1} \|\hat{g}_0(\xi, \cdot)\|_{L^2(d\gamma_k)} \\ & \quad + |\xi|^{\ell+1} |\hat{\varphi}(\xi)| \|f_0\|_{L^2(d\gamma_{k+2\ell+2}; L^1((1+|x|^\ell) dx))} \int_0^{t_0} e^{-\frac{\mu_\xi}{2}(t_0-s)} e^{-\frac{\lambda}{2}s} ds. \end{aligned}$$

We know from (10) that  $\mu_\xi = \Lambda |\xi|^2 / (1 + |\xi|^2)$  so that  $\mu_\xi \geq \frac{1}{2} \Lambda |\xi|^2$  if  $|\xi| < 1$  and  $\mu_\xi \geq \frac{1}{2} \Lambda$  if  $|\xi| \geq 1$ . Hence, for any  $t_0 \geq 1$ ,

$$\begin{aligned} & \|e^{-\frac{1}{2}\mu_\xi t_0} |\xi|^{\ell+1} \mathbf{1}_{|\xi| < 1}\|_{L^2(d\xi)} \leq \left( \int_{\mathbb{R}^d} e^{-\frac{\Lambda}{2} |\xi|^2 t_0} |\xi|^{2(\ell+1)} d\xi \right)^{\frac{1}{2}} \lesssim t_0^{-(1+\ell+\frac{d}{2})}, \\ & \int_{|\xi| \geq 1} e^{-\mu_\xi t_0} \|\hat{g}_0(\xi, \cdot)\|_{L^2(d\gamma_k)}^2 d\xi \lesssim e^{-\frac{\Lambda}{2} t_0} \|g_0\|_{L^2(dx d\gamma_k)}^2 \end{aligned}$$

by Plancherel's formula. We conclude by observing that

$$\int_{|\xi| \leq 1} |\xi|^{\ell+1} |\hat{\varphi}(\xi)| \int_0^{t_0} e^{-\frac{\mu\xi}{2}(t_0-s)} e^{-\frac{\lambda}{2}s} ds d\xi \leq \|\varphi\|_{L^1(dx)} \int_0^{t_0} \left( \int_{|\xi| \leq 1} |\xi|^{\ell+1} e^{-\frac{\Lambda}{2}|\xi|^2(t_0-s)} d\xi \right) e^{-\frac{\lambda}{2}s} ds$$

$$\lesssim t_0^{-(1+\ell+\frac{d}{2})},$$

$$\int_{|\xi| \geq 1} |\xi|^{\ell+1} |\hat{\varphi}(\xi)| \int_0^{t_0} e^{-\frac{\mu\xi}{2}(t_0-s)} e^{-\frac{\lambda}{2}s} ds d\xi \lesssim \|\xi|^{\ell+1} \hat{\varphi}(\xi)\|_{L^1(d\xi)} t_0 e^{-\frac{1}{4} \min\{\Lambda, 2\lambda\} t_0}.$$

Altogether, we obtain

$$\|g(t_0, \cdot, \cdot)\|_{L^2(dx d\gamma_k)}^2 = \|\hat{g}(t_0, \cdot, \cdot)\|_{L^2(d\xi d\gamma_k)}^2 \lesssim t_0^{-(1+\ell+\frac{d}{2})}.$$

The decay result of Theorem 3 is then obtained by writing

$$\|f_{2t_0}\|_{L^2(dx d\gamma_k)}^2 \lesssim \|g(t_0, \cdot, \cdot)\|_{L^2(dx d\gamma_k)}^2 + \left\| e^{(L-T)t_0} \left( \sum_{|\alpha| \leq \ell} \frac{1}{\alpha!} X^\alpha [f_{t_0}] \partial^\alpha \varphi \right) \right\|_{L^2(dx d\gamma_k)}$$

and using (24) for any  $t_0 \geq 1$ , with  $t = 2t_0$ . For  $t \leq 2$ , the estimate of Theorem 3 is straightforward by Corollary 8, which concludes the proof.

### Appendix A: An explicit computation of Green's function for the kinetic Fokker–Planck equation and consequences

In the whole-space case, when  $M$  is the normalized Gaussian function, let us consider the kinetic Fokker–Planck equation of Case (a)

$$\partial_t f + v \cdot \nabla_x f = \nabla_v \cdot (v f + \nabla_v f) \quad (25)$$

on  $(0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d \ni (t, x, v)$ . The characteristics associated with the equations

$$\frac{dx}{dt} = v, \quad \frac{dv}{dt} = -v$$

suggest to change variables and consider the distribution function  $g$  such that

$$f(t, x, v) = e^{dt} g(t, x + (1 - e^t)v, e^t v) \quad \text{for all } (t, x, v) \in (0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d.$$

The kinetic Fokker–Planck equation is changed into a heat equation in both variables  $x$  and  $v$  with  $t$  dependent coefficients, which can be written as

$$\partial_t g = \nabla \cdot \hat{\mathcal{D}} \nabla g, \quad (26)$$

where  $\nabla g = (\nabla_v g, \nabla_x g)$  and  $\hat{\mathcal{D}}$  is the  $t$ -derivative of the block matrix

$$\hat{\mathcal{D}} = \frac{1}{2} \begin{pmatrix} a \text{Id} & b \text{Id} \\ b \text{Id} & c \text{Id} \end{pmatrix},$$

with  $a = e^{2t} - 1$ ,  $b = 2e^t - 1 - e^{2t}$ , and  $c = e^{2t} - 4e^t + 2t + 3$ . Here  $\text{Id}$  is the identity matrix on  $\mathbb{R}^d$ . We observe that  $\mathcal{D}$  is degenerate: it is nonnegative but its lowest eigenvalue is 0. However, the change of variables allows the computation of a Green's function.

**Lemma 11.** *The Green's function of (26) is given for any  $(t, x, v) \in (0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d$  by*

$$G(t, x, v) = \frac{1}{(2\pi(ac - b^2))^{\frac{d}{2}}} \exp\left(-\frac{a|x|^2 - 2bx \cdot v + c|v|^2}{2(ac - b^2)}\right).$$

The method is standard and goes back to [Kolmogoroff 1934] (also see [Ilin and Hasminskii 1964; Hörmander 1967; Victory and O'Dwyer 1990; Bouchut 1993]).

*Proof.* By a Fourier transformation in  $x$  and  $v$ , with associated variables  $\xi$  and  $\eta$ , we find that

$$\log C - \log \widehat{G}(t, \xi, \eta) = (\eta, \xi) \cdot \mathcal{D}(\eta, \xi) = \frac{1}{2}(a|\eta|^2 + 2b\eta \cdot \xi + c|\xi|^2) = \frac{1}{2}a\left|\eta + \frac{b}{a}\xi\right|^2 + \frac{1}{2}A|\xi|^2, \quad A = c - \frac{b^2}{a},$$

for some constant  $C > 0$  which is determined by the mass normalization condition

$$\|G(t, \cdot, \cdot)\|_{L^1(\mathbb{R}^d \times \mathbb{R}^d)} = 1.$$

Let us take the inverse Fourier transform with respect to  $\eta$ ,

$$(2\pi)^{-d} \int_{\mathbb{R}^d} e^{iv \cdot \eta} \widehat{G}(t, \xi, \eta) d\eta = \frac{C}{(2\pi a)^{\frac{d}{2}}} e^{-\frac{|v|^2}{2a} - i\frac{b}{a}v \cdot \xi} e^{-\frac{1}{2}A|\xi|^2} = \frac{C}{(2\pi a)^d} e^{-\frac{|v|^2}{2a}} e^{-\frac{1}{2}A|\xi + i\frac{b}{aA}v|^2 - \frac{b^2}{2a^2A}|v|^2},$$

and then the inverse Fourier transform with respect to  $\xi$ , so that we obtain

$$G(t, x, v) = \frac{C}{(2\pi a)^{\frac{d}{2}} (2\pi A)^{\frac{d}{2}}} e^{-(1 + \frac{b^2}{aA})\frac{|v|^2}{2a}} e^{-\frac{|x|^2}{2A}} e^{\frac{b}{aA}x \cdot v} = \frac{C}{(4\pi^2 aA)^{\frac{d}{2}}} e^{-\frac{1}{2A}|x - \frac{b}{a}v|^2} e^{-\frac{|v|^2}{2a}}.$$

It is easy to check that  $C = 1$ . □

Let us consider a solution  $g$  of (26) with initial datum  $g_0 \in L^1(\mathbb{R}^d \times \mathbb{R}^d)$ . From the representation

$$g(t, \cdot, \cdot) = G(t, \cdot, \cdot) *_{x,v} g_0,$$

we obtain the estimate

$$\|g(t, \cdot, \cdot)\|_{L^\infty(\mathbb{R}^d \times \mathbb{R}^d)} \leq \|G(t, \cdot, \cdot)\|_{L^\infty(\mathbb{R}^d \times \mathbb{R}^d)} \|g_0\|_{L^1(\mathbb{R}^d \times \mathbb{R}^d)} = \frac{\|g_0\|_{L^1(\mathbb{R}^d \times \mathbb{R}^d)}}{(8\pi^2)^{\frac{d}{2}}} t^{-\frac{d}{2}} e^{-dt} (1 + O(t^{-1}))$$

as  $t \rightarrow \infty$ . As a consequence, we obtain that the solution of (25) with a nonnegative initial datum  $f_0$  satisfies

$$\|f(t, \cdot, \cdot)\|_{L^\infty(\mathbb{R}^d \times \mathbb{R}^d)} = \frac{\|f_0\|_{L^1(\mathbb{R}^d \times \mathbb{R}^d)}}{(8\pi^2 t)^{\frac{d}{2}}} (1 + o(1)) \quad \text{as } t \rightarrow \infty.$$

Using the simple Hölder interpolation inequality

$$\|f\|_{L^p(\mathbb{R}^d \times \mathbb{R}^d)} \leq \|f\|_{L^1(\mathbb{R}^d \times \mathbb{R}^d)}^{\frac{1}{p}} \|f\|_{L^\infty(\mathbb{R}^d \times \mathbb{R}^d)}^{1 - \frac{1}{p}},$$

we obtain the following decay result.

**Corollary 12.** *If  $f$  is a solution of (25) with a nonnegative initial datum  $f_0 \in L^1(\mathbb{R}^d \times \mathbb{R}^d)$ , then for any  $p \in (1, \infty]$  we have the decay estimate*

$$\|f(t, \cdot, \cdot)\|_{L^p(\mathbb{R}^d \times \mathbb{R}^d)} \leq \frac{\|f_0\|_{L^1(\mathbb{R}^d \times \mathbb{R}^d)}}{(8\pi^2 t)^{\frac{d}{2}(1-\frac{1}{p})}} (1 + o(1)) \quad \text{as } t \rightarrow \infty.$$

By taking  $f_0(x, v) = G(1, x, v)$ , it is moreover straightforward to check that this estimate is optimal. With  $p = 2$ , this also proves that the decay rate obtained in Theorem 1 for the Fokker–Planck operator, i.e., Case (a), is the optimal one because, again with  $f_0(x, v) = G(1, x, v)$ , we observe that

$$\|f(t, \cdot, \cdot)\|_{L^2(dx dy_k)}^2 = e^{dt} \|G(t, \cdot, \cdot)\|_{L^2(dx dv)}^2 = O(t^{-\frac{d}{2}}) \quad \text{as } t \rightarrow +\infty.$$

### Appendix B: Consistency with the decay rates of the heat equation

In the whole-space case, the abstract approach of [Dolbeault et al. 2015] is inspired by the diffusion limit of (1). We consider the scaled equation

$$\varepsilon \frac{dF}{dt} + \mathbb{T}F = \frac{1}{\varepsilon} \mathbb{L}F, \quad (27)$$

which formally corresponds to a parabolic rescaling given by  $t \mapsto \varepsilon^2 t$  and  $x \mapsto \varepsilon x$ , and investigate the limit as  $\varepsilon \rightarrow 0_+$ . Let us check that the rates are asymptotically independent of  $\varepsilon$  and consistent with those of the heat equation.

**B.1. Mode-by-mode hypocoercivity.** It is straightforward to check that in the estimate (7) for  $\lambda$ , the gap constant  $\lambda_m$  has to be replaced by  $\lambda_m/\varepsilon$ , while, with the notation of Proposition 4,  $C_M$  can be replaced by  $C_M/\varepsilon$  for  $\varepsilon < 1$ . In the asymptotic regime as  $\varepsilon \rightarrow 0_+$ , we obtain

$$\varepsilon \frac{d}{dt} \mathbb{H}[F] \leq -\mathbb{D}[F] \leq -\frac{\lambda_M}{3(1+\lambda_M)} \frac{\lambda_m \lambda_M \varepsilon}{(1+\lambda_M) C_M^2} \mathbb{D}[F],$$

which proves that the estimate of Proposition 4 becomes

$$\lambda \geq \frac{\lambda_m \lambda_M^2}{3(1+\lambda_M)^2 C_M^2}.$$

We observe that this rate is independent of  $\varepsilon$ .

**B.2. Decay rates based on Nash's inequality in the whole-space case.** In the proof of Theorem 1,  $\bar{\sigma}$  has to be replaced by  $\bar{\sigma}/\varepsilon$  and in the limit as  $\varepsilon \rightarrow 0_+$ , we get that  $b \sim 4\bar{\sigma}/\varepsilon$  and (16) is satisfied with

$$4a = \delta \sim \frac{\lambda_m}{8\bar{\sigma}^2} \varepsilon.$$

Hence (18) asymptotically becomes, as  $\varepsilon \rightarrow 0_+$ ,

$$-\frac{d}{dt} \mathbb{H}[f] \geq \frac{\lambda_m}{4\bar{\sigma}^2} c \left( \frac{2}{1+\delta} \mathbb{H}[f] \right)^{1+\frac{2}{a}},$$

which again gives a rate of decay which is independent of  $\varepsilon$ . The algebraic decay rate in Theorem 1 is the one of the heat equation on  $\mathbb{R}^d$  and it is independent of  $\varepsilon$  in the limit as  $\varepsilon \rightarrow 0_+$ .

**B.3. Decay rates in the whole-space case for distribution functions with moment cancellations.** The improved rate of Theorem 2 is consistent with a parabolic rescaling: if  $f$  solves (1), then  $f^\varepsilon(t, x, v) = \varepsilon^{-d} f(\varepsilon^{-2}t, \varepsilon^{-1}x, v)$  solves (27). With the notation of Section 6.1, let  $g^\varepsilon = f^\varepsilon - f_\bullet^\varepsilon \varphi(\cdot/\varepsilon)$ , with  $\varphi^\varepsilon = \varepsilon^{-d} \varphi(\cdot/\varepsilon)$ . The Fourier transform of  $g^\varepsilon$  solves

$$\varepsilon^2 \partial_t \hat{g}^\varepsilon + \varepsilon \mathbb{T} \hat{g}^\varepsilon = \mathbb{L} \hat{g}^\varepsilon - \varepsilon f_\bullet^\varepsilon \mathbb{T} \hat{\varphi}^\varepsilon.$$

The decay rate  $\lambda$  in (20) becomes  $\lambda/\varepsilon^2$  and the decay rate of the semigroup generated by  $\mathbb{L} - \varepsilon \mathbb{T}$  is, with the notation of Corollary 5,  $\mu_{\varepsilon \xi}$ . Moreover,  $\Lambda$  in (10) is given by  $\Lambda = \frac{1}{3} \min\{1, \Theta\}$  for any  $\varepsilon > 0$ , small enough. Duhamel's formula (21) has to be replaced by

$$\|\hat{g}^\varepsilon(t, \xi, \cdot)\|_{L^2(d\gamma_k)} \leq C e^{-\frac{\mu_{\varepsilon \xi}}{2\varepsilon^2}t} \|\hat{g}_0^\varepsilon(\xi, \cdot)\|_{L^2(d\gamma_k)} + C \int_0^t e^{-\frac{\mu_{\varepsilon \xi}}{2\varepsilon^2}(t-s)} \|f_\bullet^\varepsilon(s, \cdot)\|_{L^2(|v|^2 d\gamma_k)} |\varepsilon \xi| |\hat{\varphi}(\varepsilon \xi)| ds.$$

Using

$$\lim_{\varepsilon \rightarrow 0_+} \frac{\mu_{\varepsilon \xi}}{\varepsilon^2} = \lim_{\varepsilon \rightarrow 0_+} \frac{\Lambda |\xi|^2}{1 + \varepsilon^2 |\xi|^2} = \Lambda |\xi|^2,$$

a computation similar to the one of Section 6.1 shows that the first term of the right-hand side is estimated by

$$\begin{aligned} \int_{\mathbb{R}^d} e^{-\frac{\mu_{\varepsilon \xi}}{\varepsilon^2}t} \|\hat{g}_0^\varepsilon(\xi, \cdot)\|_{L^2(d\gamma_k)}^2 d\xi &= \int_{|\xi| \leq \frac{1}{\varepsilon}} e^{-\frac{\mu_{\varepsilon \xi}}{\varepsilon^2}t} \|\hat{g}_0^\varepsilon(\xi, \cdot)\|_{L^2(d\gamma_k)}^2 d\xi + \int_{|\xi| > \frac{1}{\varepsilon}} e^{-\frac{\mu_{\varepsilon \xi}}{\varepsilon^2}t} \|\hat{g}_0^\varepsilon(\xi, \cdot)\|_{L^2(d\gamma_k)}^2 d\xi \\ &\leq \|g_0^\varepsilon\|_{L^2(d\gamma_k; L^1(|x| dx))}^2 \int_{\mathbb{R}^d} |\xi|^2 e^{-\frac{\Lambda}{2} |\xi|^2 t} d\xi + \|g_0^\varepsilon\|_{L^2(dx d\gamma_k)}^2 e^{-\frac{\Lambda}{2} \frac{t}{\varepsilon^2}}, \end{aligned}$$

while the square of the second term is bounded by

$$\begin{aligned} \|f_\bullet^\varepsilon(t=0, \cdot)\|_{L^2(|v|^2 d\gamma_k)}^2 \int_{\mathbb{R}^d} |\varepsilon \xi|^2 |\hat{\varphi}(\varepsilon \xi)|^2 \left( \int_0^{\varepsilon^{-2}t} e^{-\frac{1}{2} \mu_{\varepsilon \xi} (\varepsilon^{-2}t-s)} e^{-\frac{1}{2} \lambda s} ds \right)^2 d\xi \\ \leq \|f_0\|_{L^2(|v|^2 d\gamma_k; L^1(dx))}^2 \left( C_1 \frac{\varepsilon^{d+1}}{t^{\frac{d}{2}+1}} + \frac{C_2}{\varepsilon^3} e^{-\min\{\frac{\Lambda}{2}, \lambda\} \frac{t}{\varepsilon^2}} \right). \end{aligned}$$

By collecting all terms and using Plancherel's formula, we conclude that the rate of convergence of Theorem 2 applied to the solution of (27) is independent of  $\varepsilon$ . We also notice that the scaled spatial density  $\rho_{f^\varepsilon} = \int_{\mathbb{R}^d} f^\varepsilon dv$  satisfies

$$\|\rho_{f^\varepsilon}(t, \cdot)\|_{L^2(dx)}^2 \leq \frac{\mathcal{C}_0}{(1+t)^{1+\frac{d}{2}}} \quad \text{for all } t \geq 0$$

for some positive constant  $\mathcal{C}_0$  which depends on  $f_0$  but is independent of  $\varepsilon$ . This is the decay of the heat equation with an initial datum of zero average.

Similar estimates can be obtained in the framework of Theorem 3.

### Acknowledgments

The French National Research Agency (ANR) provided partial support of Bouin and Dolbeault through Projects EFI (ANR-17-CE40-0030) and Kibord (ANR-13-BS01-0004) and of Dolbeault through Project STAB (ANR-12-BS01-0019). Schmeiser was supported by the Austrian Science Foundation (grants F65 and W1245), by the Fondation Sciences Mathématiques de Paris, and by Paris Science et Lettres. Mouhot and Bouin acknowledge partial funding by the ERC grants MATKIT 2011-2016 and MAFRAN 2017-2022. Moreover Mouhot and Schmeiser are very grateful for the hospitality at Université Paris-Dauphine.

### References

- [Achleitner et al. 2016] F. Achleitner, A. Arnold, and E. A. Carlen, “On linear hypocoercive BGK models”, pp. 1–37 in *From particle systems to partial differential equations, III*, edited by P. Gonçalves and A. J. Soares, Springer Proc. Math. Stat. **162**, Springer, 2016. MR Zbl
- [Bakry et al. 2008] D. Bakry, F. Barthe, P. Cattiaux, and A. Guillin, “A simple proof of the Poincaré inequality for a large class of probability measures including the log-concave case”, *Electron. Commun. Probab.* **13** (2008), 60–66. MR Zbl
- [Bartier et al. 2011] J.-P. Bartier, A. Blanchet, J. Dolbeault, and M. Escobedo, “Improved intermediate asymptotics for the heat equation”, *Appl. Math. Lett.* **24**:1 (2011), 76–81. MR Zbl
- [Beckner 1989] W. Beckner, “A generalized Poincaré inequality for Gaussian measures”, *Proc. Amer. Math. Soc.* **105**:2 (1989), 397–400. MR Zbl
- [Bhatnagar et al. 1954] P. L. Bhatnagar, E. P. Gross, and M. Krook, “A model for collision processes in gases, I: Small amplitude processes in charged and neutral one-component systems”, *Phys. Rev.* **94**:3 (1954), 511–525. Zbl
- [Bouchut 1993] F. Bouchut, “Existence and uniqueness of a global smooth solution for the Vlasov–Poisson–Fokker–Planck system in three dimensions”, *J. Funct. Anal.* **111**:1 (1993), 239–258. MR Zbl
- [Bouin et al. 2017] E. Bouin, F. Hoffmann, and C. Mouhot, “Exponential decay to equilibrium for a fiber lay-down process on a moving conveyor belt”, *SIAM J. Math. Anal.* **49**:4 (2017), 3233–3251. MR Zbl
- [Cáceres et al. 2003] M. J. Cáceres, J. A. Carrillo, and T. Goudon, “Equilibration rate for the linear inhomogeneous relaxation-time Boltzmann equation for charged particles”, *Comm. Partial Differential Equations* **28**:5-6 (2003), 969–989. MR Zbl
- [Degond et al. 2000] P. Degond, T. Goudon, and F. Poupaud, “Diffusion limit for nonhomogeneous and non-micro-reversible processes”, *Indiana Univ. Math. J.* **49**:3 (2000), 1175–1198. MR Zbl
- [Dolbeault et al. 2009] J. Dolbeault, C. Mouhot, and C. Schmeiser, “Hypocoercivity for kinetic equations with linear relaxation terms”, *C. R. Math. Acad. Sci. Paris* **347**:9-10 (2009), 511–516. MR Zbl
- [Dolbeault et al. 2015] J. Dolbeault, C. Mouhot, and C. Schmeiser, “Hypocoercivity for linear kinetic equations conserving mass”, *Trans. Amer. Math. Soc.* **367**:6 (2015), 3807–3828. MR Zbl
- [Duan 2011] R. Duan, “Hypocoercivity of linear degenerately dissipative kinetic equations”, *Nonlinearity* **24**:8 (2011), 2165–2189. MR Zbl
- [Eckmann and Hairer 2003] J.-P. Eckmann and M. Hairer, “Spectral properties of hypoelliptic operators”, *Comm. Math. Phys.* **235**:2 (2003), 233–253. MR Zbl
- [Evans 2017] J. Evans, “Hypocoercivity in Phi-entropy for the linear relaxation Boltzmann equation on the torus”, preprint, 2017. arXiv
- [Glasse 1996] R. T. Glassey, *The Cauchy problem in kinetic theory*, SIAM, Philadelphia, PA, 1996. MR Zbl
- [Gualdani et al. 2017] M. P. Gualdani, S. Mischler, and C. Mouhot, *Factorization of non-symmetric operators and exponential H-theorem*, Mém. Soc. Math. Fr. (N.S.) **153**, Société Mathématique de France, Paris, 2017. MR Zbl
- [Han-Kwan and Léautaud 2015] D. Han-Kwan and M. Léautaud, “Geometric analysis of the linear Boltzmann equation, I: Trend to equilibrium”, *Ann. PDE* **1**:1 (2015), art. id. 3. MR Zbl

- [Héreau 2006] F. Héreau, “Hypocoercivity and exponential time decay for the linear inhomogeneous relaxation Boltzmann equation”, *Asymptot. Anal.* **46**:3-4 (2006), 349–359. MR Zbl
- [Héreau and Nier 2004] F. Héreau and F. Nier, “Isotropic hypoellipticity and trend to equilibrium for the Fokker–Planck equation with a high-degree potential”, *Arch. Ration. Mech. Anal.* **171**:2 (2004), 151–218. MR Zbl
- [Hörmander 1967] L. Hörmander, “Hypoelliptic second order differential equations”, *Acta Math.* **119** (1967), 147–171. MR Zbl
- [Iacobucci et al. 2019] A. Iacobucci, S. Olla, and G. Stoltz, “Convergence rates for nonequilibrium Langevin dynamics”, *Ann. Math. Qué.* **43**:1 (2019), 73–98. MR Zbl
- [Ilin and Hasminskii 1964] A. M. Ilin and R. Z. Hasminskii, “On the equations of Brownian motion”, *Teor. Veroyatnost. i Primenen.* **9** (1964), 466–491. In Russian; translated in *Th. Prob. Appl.* **9**:3 (1964), 421–444. MR Zbl
- [Kavian and Mischler 2015] O. Kavian and S. Mischler, “The Fokker–Planck equation with subcritical confinement force”, preprint, 2015. arXiv
- [Kawashima 1990] S. Kawashima, “The Boltzmann equation and thirteen moments”, *Japan J. Appl. Math.* **7**:2 (1990), 301–320. MR Zbl
- [Kolmogoroff 1934] A. Kolmogoroff, “Zufällige Bewegungen (zur Theorie der Brownschen Bewegung)”, *Ann. of Math.* (2) **35**:1 (1934), 116–117. MR Zbl
- [Mischler and Mouhot 2016] S. Mischler and C. Mouhot, “Exponential stability of slowly decaying solutions to the kinetic-Fokker–Planck equation”, *Arch. Ration. Mech. Anal.* **221**:2 (2016), 677–723. MR Zbl
- [Monmarché 2017] P. Monmarché, “A note on Fisher information hypocoercive decay for the linear Boltzmann equation”, preprint, 2017. arXiv
- [Mouhot and Neumann 2006] C. Mouhot and L. Neumann, “Quantitative perturbative study of convergence to equilibrium for collisional kinetic models in the torus”, *Nonlinearity* **19**:4 (2006), 969–998. MR Zbl
- [Nash 1958] J. Nash, “Continuity of solutions of parabolic and elliptic equations”, *Amer. J. Math.* **80** (1958), 931–954. MR Zbl
- [Reed and Simon 1980] M. Reed and B. Simon, *Methods of modern mathematical physics, I: Functional analysis*, 2nd ed., Academic Press, New York, 1980. MR Zbl
- [Vázquez 2017] J. L. Vázquez, “Asymptotic behaviour methods for the heat equation: convergence to the Gaussian”, preprint, 2017. arXiv
- [Victory and O’Dwyer 1990] H. D. Victory, Jr. and B. P. O’Dwyer, “On classical solutions of Vlasov–Poisson Fokker–Planck systems”, *Indiana Univ. Math. J.* **39**:1 (1990), 105–156. MR Zbl
- [Villani 2006] C. Villani, “Hypocoercive diffusion operators”, pp. 473–498 in *International Congress of Mathematicians*, vol. III, edited by M. Sanz-Solé et al., Eur. Math. Soc., Zürich, 2006. MR Zbl
- [Villani 2009] C. Villani, *Hypocoercivity*, Mem. Amer. Math. Soc. **950**, Amer. Math. Soc., Providence, RI, 2009. MR Zbl

Received 7 Nov 2018. Revised 20 Sep 2019. Accepted 24 Nov 2019.

EMERIC BOUIN: bouin@ceremade.dauphine.fr

Ceremade (CNRS UMR 7534), PSL University, Université Paris-Dauphine, Paris, France

JEAN DOLBEAULT: dolbeaul@ceremade.dauphine.fr

Ceremade (CNRS UMR 7534), PSL University, Université Paris-Dauphine, Paris, France

STÉPHANE MISCHLER: mischler@ceremade.dauphine.fr

Ceremade (CNRS UMR 7534), PSL University, Université Paris-Dauphine, Paris, France

CLÉMENT MOUHOT: c.mouhot@dpms.cam.ac.uk

DPMS, Centre for Mathematical Sciences, University of Cambridge, Cambridge, United Kingdom

CHRISTIAN SCHMEISER: christian.schmeiser@univie.ac.at

Fakultät für Mathematik, Universität Wien, Wien, Austria

# PURE and APPLIED ANALYSIS

msp.org/paa

## EDITORS-IN-CHIEF

- Charles L. Epstein University of Pennsylvania  
cle@math.upenn.edu
- Maciej Zworski University of California at Berkeley  
zworski@math.berkeley.edu

## EDITORIAL BOARD

- Sir John M. Ball University of Oxford  
ball@maths.ox.ac.uk
- Michael P. Brenner Harvard University  
brenner@seas.harvard.edu
- Charles Fefferman Princeton University  
cf@math.princeton.edu
- Susan Friedlander University of Southern California  
susanfri@usc.edu
- Anna Gilbert University of Michigan  
annacg@umich.edu
- Leslie F. Greengard Courant Institute, New York University, and  
Flatiron Institute, Simons Foundation  
greengard@cims.nyu.edu
- Yan Guo Brown University  
yan\_guo@brown.edu
- Claude Le Bris CERMICS - ENPC  
lebris@cermics.enpc.fr
- Robert J. McCann University of Toronto  
mccann@math.toronto.edu
- Michael O'Neil Courant Institute, New York University  
oneil@cims.nyu.edu
- Galina Perelman Université Paris-Est Créteil  
galina.perelman@u-pec.fr
- Jill Pipher Brown University  
jill\_pipher@brown.edu
- Johannes Sjöstrand Université de Dijon  
johannes.sjostrand@u-bourgogne.fr
- Vladimir Šverák University of Minnesota  
sverak@math.umn.edu
- Daniel Tataru University of California at Berkeley  
tataru@berkeley.edu
- Michael I. Weinstein Columbia University  
miw2103@columbia.edu
- Jon Wilkening University of California at Berkeley  
wilken@math.berkeley.edu
- Enrique Zuazua DeustoTech-Bilbao, and  
Universidad Autónoma de Madrid  
enrique.zuazua@deusto.es

## PRODUCTION

- Silvio Levy (Scientific Editor)  
production@msp.org

**Cover image:** The figure shows the outgoing scattered field produced by scattering a plane wave, coming from the northwest, off of the (stylized) letters P A A. The total field satisfies the homogeneous Dirichlet condition on the boundary of the letters. It is based on a numerical computation by Mike O'Neil of the Courant Institute.

---

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

---

The subscription price for 2020 is US \$505/year for the electronic version, and \$565/year (+\$25, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

---

Pure and Applied Analysis (ISSN 2578-5885 electronic, 2578-5893 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

---

PAA peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

 **mathematical sciences publishers**  
nonprofit scientific publishing  
<http://msp.org/>

© 2020 Mathematical Sciences Publishers

# PURE and APPLIED ANALYSIS

vol. 2 no. 2 2020

Hypocoercivity without confinement	203
EMERIC BOUIN, JEAN DOLBEAULT, STÉPHANE MISCHLER, CLÉMENT MOUHOT and CHRISTIAN SCHMEISER	
The hyperbolic Yang–Mills equation in the caloric gauge: local well-posedness and control of energy-dispersed solutions	233
SUNG-JIN OH and DANIEL TATARU	
Scattering resonances on truncated cones	385
DEAN BASKIN and MENGXUAN YANG	
The Kähler geometry of certain optimal transport problems	397
GABRIEL KHAN and JUN ZHANG	
Semiclassical asymptotics for nonselfadjoint harmonic oscillators	427
VÍCTOR ARNAIZ and GABRIEL RIVIÈRE	
On the solution of Laplace’s equation in the vicinity of triple junctions	447
JEREMY HOSKINS and MANAS RACHH	
Enhanced convergence rates and asymptotics for a dispersive Boussinesq-type system with large ill-prepared data	477
FRÉDÉRIC CHARVE	
Dispersive estimates, blow-up and failure of Strichartz estimates for the Schrödinger equation with slowly decaying initial data	519
RAINER MANDEL	



2578-5893(2020)2:2;1-N