Exponential convergence to equilibrium for kinetic Fokker-Planck equations on Riemannian manifolds

Simone Calogero

Departamento de Matemática Aplicada
Universidad de Granada, Spain
E-Mail: calogero@ugr.es

Abstract

A class of linear kinetic Fokker-Planck equations with a non-trivial diffusion matrix and with periodic boundary conditions in the spatial variable is considered. After formulating the problem in a geometric setting, the question of the rate of convergence to equilibrium is studied within the formalism of differential calculus on Riemannian manifolds. Under explicit geometric assumptions on the velocity field, the energy function and the diffusion matrix, it is shown that global regular solutions converge in time to equilibrium with exponential rate. The result is proved by estimating the time derivative of a modified entropy functional, as recently proposed by Villani. For spatially homogeneous solutions the assumptions of the main theorem reduce to the curvature bound condition for the validity of logarithmic Sobolev inequalities discovered by Bakry and Emery.

1 Introduction

The evolution of many physical or biological systems is characterized by two kinds of driving mechanism: diffusion and friction. The competition between these two types of dynamics may lead the system to a thermodynamical equilibrium. The purpose of this paper is to study the rate of convergence to equilibrium for a class of linear models that exhibit this kind of behavior. The simplest model in this class is the Fokker-Planck equation [13] on the density function of the system:

$$\partial_t \rho = \Delta\rho + \nabla\cdot(\xi\rho)\,, \qquad t>0\,, \quad \xi\in\mathbb{R}^N\,. \tag{1}$$

In this model, the density function ρ depends on the variables (t, ξ), the diffusion term is given by $\Delta\rho$ and the friction term by $\nabla \cdot (\xi\rho)$. All sufficiently regular solutions of (1) converge in time to a Maxwellian type distribution, with exponential rate of convergence [4]. The convergence holds for instance in the L^1-norm.

An important generalization of (1), often considered in the mathematical and in the physical literature [1, 13], is

$$\partial_t \rho = \nabla \cdot (D(\nabla\rho + \rho\nabla E)) \,, \tag{2}$$

where $D = D(\xi)$ is the diffusion matrix and $E = E(\xi)$ is the energy function ($E = |\xi|^2/2$ for (1)). It is assumed that D is positive definite and that

$$\Theta^{-1} = \int_{\mathbb{R}^N} e^{-E} \, d\xi$$

is bounded. Equation (2) admits an unique invariant probability measure, which is given by $d\mu = \rho_\infty(\xi)\, d\xi$, where $\rho_\infty = \Theta\, e^{-E}$. Non-negative solutions of (2) with unit mass converge to the equilibrium state ρ_∞ with exponential rate if the matrix D and the function E satisfy an inequality known as the curvature bound condition [2, 3]. Let us briefly recall the argument of the proof. In terms of $h(t,\xi) = \rho(t,\xi)/\rho_\infty(\xi)$, equation (2) takes the form

$$\partial_t h = \nabla \cdot (D\nabla h) - D\nabla E \cdot \nabla h \,,$$

or, equivalently,

$$\partial_t h = \Delta^G h + Qh \,, \tag{3}$$

where Δ^G denotes the Laplace-Beltrami operator associated to the Riemannian metric $G = D^{-1}$ and Q is the vector field

$$Qh = D\nabla \log u \cdot \nabla h \,, \quad u = \sqrt{\det D}\, e^{-E} \,.$$

The entropy functional associated to (3) is given by

$$\mathfrak{D}[h] = \int_{\mathbb{R}^N} h \log h \, d\mu \,, \tag{4}$$

and satisfies

$$\frac{d}{dt}\mathfrak{D}[h] = -\mathfrak{I}[h] \,, \quad \text{where} \;\; \mathfrak{I}[h] = \int_{\mathbb{R}^N} \frac{D\nabla h \cdot \nabla h}{h} \, d\mu \tag{5}$$

is the entropy dissipation functional. Let Ric^G and ∇^G denote the Ricci curvature and the Levi-Civita connection of G. Bakry and Emery proved in [2, 3] that if the curvature bound condition

$$\mathrm{Ric}^G - \nabla^G Q \geq \alpha\, G \,, \quad \text{for some} \;\; \alpha > 0 \,, \tag{6}$$

is satisfied, then the logarithmic Sobolev inequality $\mathfrak{D}[f] \leq (2\alpha)^{-1}\mathfrak{I}[f]$ holds for all sufficiently regular probability densities f. Replacing in (5) we obtain

$$\frac{d}{dt}\mathfrak{D}[h] \leq -2\alpha\,\mathfrak{D}[h] \,,$$

whence the entropy functional decays exponentially as $O(e^{-2\alpha t})$. The classical Csiszár-Kullback inequality [6],

$$\|h-1\|_{L^1(d\mu)} \leq \sqrt{2\mathfrak{D}} \,, \tag{7}$$

implies that h converges to 1 as $t \to \infty$ in $L^1(d\mu)$ with exponential rate or, equivalently, the solution ρ of (2) converges to ρ_∞ in $L^1(d\xi)$ with exponential rate.

Fokker-Planck type equations appear also in kinetic theory and these are the subject of the present investigation. Assuming periodic boundary conditions in space, the simplest kinetic Fokker-Planck (or Kramers) equation is given by [13]

$$\partial_t f + p \cdot \nabla_x f = \Delta_p f + \nabla_p \cdot (pf) \,, \quad t > 0 \,,\ x \in \mathbb{T}^N \,,\ p \in \mathbb{R}^N \,. \tag{8}$$

Here $f = f(t,x,p) \geq 0$ is the particles distribution in phase-space, with (t,x) denoting the space-time variables and p the momentum variable. Equation (8) describes the kinetic motion of a system of particles undergoing stochastic collisions with the molecules of a homogeneous fluid in thermal equilibrium. From a mathematical point of view, (8) is more complicated than (1), due to the presence of the transport derivative $p \cdot \nabla_x$ and the fact that the diffusion operator Δ_p is degenerate, i.e., it acts only on the momentum variables. Moreover equation (1) can be seen as the spatially homogeneous version of (8).

The problem of how fast the solutions of (8) converge to equilibrium was solved only quite recently by Hérau and Nier [9] and by Villani [14]. Both references establish exponential convergence to equilibrium, however by completely different methods. In [9] the problem is tackled by spectral analysis techniques—exponential rate of convergence is implied by the existence of a spectral gap in the spectrum of the Fokker-Planck operator—, while the proof given in [14, Th. 28] is based on the study of the evolution of a properly modified entropy functional. (See [7] for an earlier study of the trend to equilibrium for (8). In the latter reference the authors prove that convergence to equilibrium occurs as fast as $O(t^{-1/\varepsilon})$, for any $\varepsilon > 0$.)

In this paper the entropy method is applied to study the trend to equilibrium for the following generalization of (8):

$$\partial_t f + v(p) \cdot \nabla_x f = \nabla_p(D(\nabla_p f + f\nabla_p E)) \,, \tag{9}$$

where $D = D(p)$ is the diffusion matrix and $E = E(p)$ is the energy function. Equation (9) reduces to (2) in the spatially homogeneous case. The vector field v in the transport term is the velocity field; the two most important examples for the applications are the classical velocity field

$$v(p) = p$$

and the relativistic velocity field

$$v(p) = \frac{p}{\sqrt{1+|p|^2}} \, .$$

Setting $h = f/e^{-E}$, we may rewrite (9) in the form

$$\partial_t h + v(p)\cdot\nabla_x h = \Delta_p h + Wh \, , \quad t>0 \, , \; x\in\mathbb{T}^N \, , \; p\in\mathbb{R}^N \, , \tag{10}$$

where Δ_p is the Laplace-Beltrami operator associated to the metric $g = D^{-1}$ and

$$Wh = D\nabla_p \log u \cdot \nabla_p h \, , \quad u = \sqrt{\det D}\, e^{-E} \, .$$

The main result of this paper is presented in Section 3. It is proved that under suitable assumptions on the functions v, D, E , which take the form of geometric inequalities involving v, g and W, smooth solutions of (10) with unit mass converge in time to the equilibrium state $h_\infty \equiv 1$ with exponential rate of convergence. The proof follows the argument previously outlined to establish exponential convergence for spatially homogeneous solutions, the main difference being that the entropy functional (4) is replaced by a properly modified entropy functional, as proposed recently by Villani [14].

It should be noticed that the trend to equilibrium for (9) is also studied in [14]. The main differences between Villani's approach and the one used in this paper are the following. In [14] the author exploits the fact that (9) can be written in the form

$$\partial_t h = (A^*A + B)h \, ,$$

where $B = -v(p)\cdot\nabla_x$, A^* is the adjoint of A in the Hilbert space $L^2(d\mu)$ and $A = \sigma\nabla$, $\sigma = \sqrt{D}$. The proof of [14, Th. 28] makes crucial use of the iterated commutators

$$[A,B] \, , \quad [B,[A,B]] \, , \quad [B,[B,[A,B]]] \, , \quad \dots$$

in the spirit of Hörmander's hypoellipticity theory [10]. However the use of commutators, while natural in the context of regularity theory, presents some disadvantages for the problem of convergence to equilibrium. In particular it leads to very heavy and sometimes obscure calculations, which, as pointed out by Villani at the beginning of the proof of Lemma 32 in [14], "might be an indication that a more appropriate formalism is still to be found". The main theorem of the present paper is proved using the formalism of differential geometry, which helps to clarify the meaning of many long expressions that have to be controlled in Villani's work. Moreover, given a diffusion matrix D, a velocity field v and an energy function E, one may check directly if our assumptions are verified. As opposed to this, one has to find a suitable way to decompose the vector fields A, B and the iterated commutators in order to verify the assumptions of [14, Th. 28]. Apart from the different approach to the problem, several ideas introduced by Villani in his important work will be adapted (and simplified) to the present context, resulting in a less technical and more elegant proof.

Finally, we remark that a natural and interesting extension of the results of this paper would be to replace $x \in \mathbb{T}^N$ with $x \in \mathbb{R}^N$ and to introduce an external confining potential in the equation. This generalization will be considered in a separate paper.

2 Set-up

This section is devoted to introduce the transport-diffusive equation that will be the subject of our study, as well as some geometric tools that are needed for this purpose.

Let $\mathcal{N}$ be a N-dimensional smooth manifold and $\mathcal{M}$ a smooth M-dimensional manifold with a C^2 Riemannian metric g. It is assumed that $\mathcal{N}$ and $\mathcal{M}$ are globally homeomorphic to the torus $\mathbb{T}^N$ and to $\mathbb{R}^M$, respectively. Let

$$h : \mathbb{R} \times \mathcal{N} \times \mathcal{M} \to [0, \infty)$$

satisfy an evolution equation of the following form:

$$\partial_t h + Th = \Delta_p h + Wh \,. \tag{11}$$

Here $h = h(t, x, p)$, where $x = (x^1, \ldots, x^N)$, $p = (p^1, \ldots, p^M)$ are global coordinates on $\mathcal{N}$ and $\mathcal{M}$ respectively; ∂_t denotes the partial derivative with respect to $t \in \mathbb{R}$, while ∂_{x^I}, ∂_{p^i} denote the partial derivatives in the coordinates (x^I, p^i). Capital Latin indexes run from 1 to N, small Latin indexes run from 1 to M. Denoting $\mathfrak{X}(\mathcal{B})$, $\mathfrak{X}_*(\mathcal{B})$ the set of smooth vector fields and one form fields on a manifold $\mathcal{B}$, then $T \in \mathfrak{X}(\mathcal{N} \times \mathcal{M})$, the *transport field*, whereas $W \in \mathfrak{X}(\mathcal{M})$. Finally, Δ_p denotes the Laplace-Beltrami operator on $(\mathcal{M}, g)$. A subscript p is attached to differential operators that act on the variables $p^1, \ldots, p^M$ only.

In order to specify the exact form of the fields T, W, some basic facts from differential geometry are required. In the following discussion, which is based mainly on [12], we consider only (smooth, time dependent) tensor fields defined on $\mathcal{M}$ or $\mathcal{N}$, possibly obtained by projecting tensor fields from $\mathcal{N} \times \mathcal{M}$. We also remark that in the rest of the paper we do not distinguish between a tensor field defined on $\mathcal{N}$ or $\mathcal{M}$ and its lift on $\mathcal{N} \times \mathcal{M}$. Let $X_{(I)}$, $P_{(i)}$ denote the frame vector fields basis associated to the coordinates x^I, p^i (i.e., $X_{(I)} f = \partial_{x^I} f$, $P_{(i)} f = \partial_{p^i} f$, for all smooth functions f on $\mathcal{M}$) and $X_*^{(I)}, P_*^{(i)}$ their dual one form fields; clearly $P_{(i)}$ and $P_*^{(i)}$ are metrically equivalent: $g(P_{(i)}, Y) = P_*^{(i)}(Y)$, for all $Y \in \mathfrak{X}(\mathcal{M})$. Note that the indexes in round brackets are list indexes and not component indexes, that is to say, for each fixed i, $P_{(i)}$ is a geometric object of the same type (a vector field). Now let $v^{(1)}, \cdots v^{(M)}$ denote a set of C^3 real valued functions on $\mathcal{M}$. We assume that the transport vector field has the following form:

$$T = v^{(I)}(p) X_{(I)} \,.$$

We adopt the Einstein summation rule, whereby the sum over repeated index is understood.

Remark. When $N = M$ (or more generally when $N \leq M$) the functions $v^{(I)}$ can be thought of as the (non-zero) components of a vector field over $\mathcal{M}$, the *velocity field*. However this interpretation is not necessary and in general not very useful, so we will refrain from adopting it. In particular, the use of the list index (I) in $v^{(I)}$ reminds that this is a scalar function, which affects how geometric differential operators act on it.

For any tensor field R over $\mathcal{M}$, $\nabla_p R$ denotes the covariant differential of R, where ∇_p is the Levi-Civita connection associated to g (i.e.,∇_p is symmetric and $\nabla_p\, g = 0$). For a scalar function f on $\mathcal{M}$, $\nabla_p f$ is the one form $\nabla_p f(Y) = Y(f)$, for all $Y \in \mathfrak{X}(\mathcal{M})$. Any vector field $Z \in \mathfrak{X}(\mathcal{M})$ is metrically equivalent to the one form field $Z_* \in \mathfrak{X}_*(\mathcal{M})$ given by $Z_*(Y) = g(Y, Z)$, for all $Y \in \mathfrak{X}(\mathcal{M})$. The vector field metrically equivalent to $\nabla_p f$ is the gradient of f, which we denote $\partial_p f$:

$$g(\partial_p f, Y) = \nabla_p f(Y)\ , \ \text{ or } \ \nabla_p f = (\partial_p f)_*\ .$$

Using the components $g_{ij} = g(P_{(i)}, P_{(j)})$ of the metric in the base $P_*^{(i)} \otimes P_*^{(j)}$ of the space of type (2, 0) tensor fields, we may express the action of the Laplace-Beltrami operator on scalar functions as

$$\Delta_p f = \frac{1}{\sqrt{|g|}} \partial_{p^i} \left(\sqrt{|g|}\, g^{ij} \partial_{p^j} f \right)\ , \tag{12}$$

where g^{ij} is the matrix inverse of g_{ij}, i.e., $g^{ik} g_{kj} = \delta^i{}_j$ and $|g| = \det g$. There is however a more convenient way to express $\Delta_p f$. For this we recall that the divergence of a vector field $Z \in \mathfrak{X}(\mathcal{M})$ is the contraction of $\nabla_p Z$, i.e.,

$$\mathrm{div}_p\, Z = \nabla_p Z(P_*^{(i)}, P_{(i)})\ .$$

We have the well known formula

$$\Delta_p f = \mathrm{div}_p\, (\partial_p f)\ .$$

Moreover by Stokes theorem

$$\int_{\mathbb{R}^M} \mathrm{div}_p\, Z \sqrt{|g|}\, dp = 0\ , \tag{13}$$

for all vector fields $Z \in H^1(\mathbb{R}^M, \sqrt{|g|}\, dp)$.

Next the definitions of the gradient of a vector field and of the divergence of a second order tensor will be recalled. Let $\nabla_p Z : \mathfrak{X}_*(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathbb{R}$ be the covariant differential of $Z \in \mathfrak{X}(\mathcal{M})$. The metrically equivalent type (2, 0) tensor field $\partial_p Z : \mathfrak{X}_*(\mathcal{M}) \times \mathfrak{X}_*(\mathcal{M}) \to \mathbb{R}$ given by

$$\partial_p Z(X_*, Y_*) = \nabla_p Z(X_*, Y)$$

is called the gradient of Z. Moreover, given any type (2, 0) tensor field R, its divergence is defined as the contraction of $\nabla_p R$ in the second and third variable, i.e.,

$$\operatorname{div}_p R(Y_*) = \nabla_p R(Y_*, P_*^{(i)}, P_{(i)})$$

and thus it is a vector field on $\mathcal{M}$. The following lemma collects some useful identities on the geometric objects defined above.

Lemma 1. *Let f, f_1, f_2 be smooth real valued functions on $\mathcal{M}$ and $X, Y, Z \in \mathfrak{X}(\mathcal{M})$. Then*

(i) $\partial_p(f_1 f_2) = f_1 \partial_p f_2 + f_1 \partial_p f_2$;

(ii) $\operatorname{div}_p(fZ) = g(\partial_p f, Z) + f \operatorname{div}_p Z$;

(iii) $\operatorname{div}_p(\partial_p f_1 \otimes \partial_p f_2)(Y_*) = \partial_p^2 f_1(\nabla_p f_2, Y_*) + \Delta_p f_2\, \partial_p f_1(Y_*)$;

(iv) $g(X, \partial_p(g(Y, Z))) = \partial_p Y(Z_*, X_*) + \partial_p Z(Y_*, X_*)$;

(v) $g(Y, \operatorname{div}_p \partial_p^2 f) = g(Y, \partial_p(\Delta_p f)) + \operatorname{Ric}(Y, \partial_p f)$,

where Ric *denotes the Ricci curvature tensor of g.*

Proof. The proofs of (i)–(iii) follow by the chain rule. The identity (iv) is a consequence of Koszul's formula applied to the Levi-Civita connection; the proof can be found in [12, Ch. 3, Th. 11]. The identity (v) is a direct consequence of the definition of the Riemann tensor and is proved for instance in [5, Lemma 1.45]. □

We can now define the vector field W. Let $E : \mathbb{R}^M \to [0, \infty)$, $E \in C^2$, such that

$$\Theta^{-1} = \int_{\mathbb{R}^M} e^{-E}\, dp$$

is bounded. Let $d\mu = \Theta e^{-E} dp$, a probability measure on $\mathbb{R}^M$. Then

$$d\mu = \Theta\, u \sqrt{|g|}\, dp\,, \quad \text{where} \quad u = \frac{e^{-E}}{\sqrt{|g|}}\,. \tag{14}$$

Let L be the operator in the r.h.s. of (11), that is

$$Lh = \Delta_p h + Wh\,.$$

We require the field W to be such that L is symmetric in the Hilbert space $L^2(d\mu) := L^2(\mathbb{R}^M, d\mu)$, i.e.,

$$\int_{\mathbb{R}^M} hLf\, d\mu = \int_{\mathbb{R}^M} fLh\, d\mu\,. \tag{15}$$

Lemma 2. *The identity* (15) *is verified if and only if* $W = \partial_p \log u$*, or equivalently,* $W_* = \nabla_p \log u$*.*

Proof. We have

$$\int_{\mathbb{R}^M} hLf\, d\mu = \Theta \int_{\mathbb{R}^M} h\,(\Delta_p f)\, u\,\sqrt{|g|}\, dp + \int_{\mathbb{R}^M} hWf\, d\mu\,.$$

In the previous equation we use (i)-(ii) of Lemma 1 to get

$$\mathrm{div}_p\,(h\,u\,\partial_p f) = h\,u\,\Delta_p f + h\,g(\partial_p f, \partial_p u) + u\,g(\partial_p f, \partial_p h)$$

and so doing we obtain, by (13),

$$\int_{\mathbb{R}^M} hLf\, d\mu = \int_{\mathbb{R}^M} h(Wf - g(\partial_p f, \partial_p \log u))\, d\mu - \int_{\mathbb{R}^M} g(\partial_p f, \partial_p h)\, d\mu\,. \tag{16}$$

Again we use

$$\mathrm{div}_p\,(f\,u\,\partial_p h) = f\,u\,\Delta_p h + f\,g(\partial_p h, \partial_p u) + u\,g(\partial_p f, \partial_p h)$$

and so we obtain

$$\int_{\mathbb{R}^M} hLf\, d\mu = \int_{\mathbb{R}^N} h(W - \partial_p \log u)f\, d\mu + \int_{\mathbb{R}^M} f(\Delta_p + \partial_p \log u)h\, d\mu\,,$$

which implies the claim. □

We conclude this section by proving some integration by parts formulas.

Lemma 3. *The following identities hold true, for all smooth real valued functions* f, h *on* $\mathcal{M}$*:*

$$\int_{\mathbb{T}^N} hTf\, dx = -\int_{\mathbb{T}^N} fTh\, dx\,, \tag{17}$$

$$\int_{\mathbb{R}^M} hLf\, d\mu = -\int_{\mathbb{R}^M} g(\partial_p f, \partial_p h)\, d\mu\,. \tag{18}$$

Proof. The proof of (17) is straightforward. The identity (18) follows by setting $W = \partial_p \log u$ in (16). □

For the next result we need to recall the definition of inner product of second order tensor fields. Given a type $(2,0)$ tensor field R and a type $(0,2)$ tensor field S, the inner product $R \cdot S = S \cdot R$ is defined as

$$R \cdot S = (R \otimes S)(P_*^{(i)}, P_*^{(j)}, P_{(i)}, P_{(j)})\,.$$

Componentwise this means $R \cdot S = R^{ij} S_{ij}$.

Lemma 4. *For all type* $(2,0)$ *tensor fields* A *and* $Z \in \mathfrak{X}(\mathcal{M})$ *we have*

$$\int_{\mathbb{R}^M} g(Z, \mathrm{div}_p\, A)\, d\mu = -\int_{\mathbb{R}^M} A \cdot \nabla_p Z_*\, d\mu - \int_{\mathbb{R}^M} A(W_*, Z_*)\, d\mu\,. \tag{19}$$

Proof. Consider the vector field $Y \in \mathfrak{X}(\mathcal{M})$ defined by $Y(\cdot) = A(uZ_*, \cdot)$. By the chain rule

$$\mathrm{div}_p\, Y = g(uZ, \mathrm{div}_p\, A) + A \cdot \nabla_p(uZ_*)\,.$$

Replacing in the l.h.s. of (19) we obtain

$$\begin{aligned}\int_{\mathbb{R}^M} g(Z, \mathrm{div}_p\, A)\, d\mu &= \Theta \int_{\mathbb{R}^M} g(uZ, \mathrm{div}_p\, A)\sqrt{|g|}\, dp \\ &= -\,\Theta \int_{\mathbb{R}^M} A \cdot \nabla_p(uZ_*)\sqrt{|g|}\, dp \\ &= -\int_{\mathbb{R}^M} A \cdot W_* \otimes Z_*\, d\mu - \int_{\mathbb{R}^M} A \cdot \nabla_p Z_*\, d\mu\,,\end{aligned}$$

which is the claim. □

3 Main result

We begin by stating our assumptions on the functions $v^{(I)}, E$ and the metric g. Let us recall that the Bakry-Emery-Ricci tensor is defined by

$$\widetilde{\mathrm{Ric}} = \mathrm{Ric} - \nabla_p W_* = \mathrm{Ric} - \nabla_p^2 \log u\,, \tag{20}$$

where $\nabla_p^2 f$ denotes the Hessian of f and u is the function (14). As already mentioned in the Introduction, Barky and Emery proved in [2, 3] that spatially homogeneous solutions of (11) converge exponentially fast in time to the equilibrium state $h_\infty \equiv 1$ in the entropic sense (i.e., the entropy functional decays exponentially) if the tensor $\widetilde{\mathrm{Ric}}$ is bounded below by a constant times the metric g. In the spatially inhomogeneous case, we also need a bound on $\widetilde{\mathrm{Ric}}$ from above.

Assumption 1. *There exist two constants* $\sigma_2 \geq \sigma_1 \geq 0$ *such that*

$$\sigma_1 g(X,X) \leq \widetilde{\mathrm{Ric}}(X,X) \leq \sigma_2 g(X,X)\,, \qquad \textit{for all } X \in \mathfrak{X}(\mathcal{M})\,.$$

We denote

$$\sigma = \sigma_2 - \sigma_1 \geq 0\,. \tag{21}$$

Before stating the next assumption, it is convenient to give the following definition.

Definition 1. *Given a real valued function f on $\mathcal{N} \times \mathcal{M}$, and a point $x \in \mathcal{N}$, we denote $\mathcal{A}_x f$ the vector field over $\mathcal{M}$ given by*

$$\mathcal{A}_x f = (\partial_{x^I} f) \partial_p v^{(I)} \,, \quad \text{evaluated at } x \in \mathcal{N}.$$

The metrically equivalent one form field is given by $(\mathcal{A}_x f)_* = (\partial_{x^I} f) \nabla_p v^{(I)}$.

Remark. We emphasize that $\mathcal{A}_x f \in \mathfrak{X}(\mathcal{M})$. (More precisely, $\mathcal{A}_x f \in \mathfrak{X}(\{x\} \times \mathcal{M}) \simeq \mathfrak{X}(\mathcal{M})$.) Its components in the vector fields basis $P_{(i)}$ are given by

$$(\mathcal{A}_x f)^i = g^{ij} \partial_{p^j} v^{(I)} \partial_{x^I} f \,.$$

For the Fokker-Planck equation (9), the manifold $\mathcal{M}$ can be identified with the tangent space at all points $x \in \mathcal{N}$ and $\mathcal{A}_x f$ coincides with $\partial_x f$, the gradient in x of f.

Now let us define a symmetric bilinear form A on $\mathbb{R}^N \times \mathbb{R}^N$ by

$$A(\xi, \eta) = A^{IJ} \xi_I \eta_J \,, \quad A^{IJ} = g(\partial_p v^{(I)}, \partial_p v^{(J)}) \,, \qquad \xi, \eta \in \mathbb{R}^N \,.$$

Note that

$$A^{IJ} \partial_{x^I} h \partial_{x^J} h = g(\mathcal{A}_x h, \mathcal{A}_x h) \,. \tag{22}$$

Assumption 2. *We assume that A is positive definite,*

$$A(\xi, \xi) > 0 \,, \quad \text{for all } 0 \neq \xi \in \mathbb{R}^N \,.$$

The next assumption is most conveniently expressed in terms of an auxiliary metric G on $\mathcal{M} \times \mathcal{N}$ and the vector field $Q \in \mathfrak{X}(\mathcal{M})$ defined as follows.

Definition 2. *Let A_{IJ} denote the matrix inverse of A^{IJ}, i.e., $A^{IJ} A_{JK} = \delta^I_K$. We define the Riemannian metric G on $\mathcal{M} \times \mathcal{N}$ as*

$$G = g_{ij} dp^i \otimes dp^j + A_{IJ} dx^I \otimes dx^J \,. \tag{23}$$

Moreover we define the vector field $Q \in \mathfrak{X}(\mathcal{M})$ as

$$Q = W - \partial_p \log \sqrt{\det(A_{IJ})} \,. \tag{24}$$

Assumption 3. *We assume that there exists a constant $\alpha \geq 0$ such that*

$$\mathrm{Ric}^G(Z, Z) - (\nabla^G Q_*)(Z, Z) \geq \alpha \, G(Z, Z) \,, \quad \text{for all } Z \in \mathfrak{X}(\mathcal{M} \times \mathcal{N}) \,,$$

where Ric^G is the Ricci tensor of G and ∇^G is the covariant differential associated to G.

Remark. Using the relations between Ric^G, ∇^G and Ric, ∇_p, Assumption 3 can be expressed in terms of inequalities on the quantities g, W. These inequalities are in general very complicated, unless A_{IJ} enjoys some simple structure, as in the statement of Corollary 1 below.

The previous assumptions suffice if the metric g and the velocity field v are such that $\nabla_p^2 v^{(I)} = 0$. If this is not the case we need more assumptions, which we give after the following definitions.

Definition 3. *Given a real valued function f on $\mathcal{N} \times \mathcal{M}$, and a point $x \in \mathcal{N}$, we denote $\mathcal{B}_x f$ the vector field over $\mathcal{M}$ given by*

$$\mathcal{B}_x f = (\partial_{x^I} f) \mathrm{div}_p \, \partial_p^2 v^{(I)} , \quad \textit{evaluated at } x \in \mathcal{N}.$$

Definition 4. *Given a real valued function f on $\mathcal{N} \times \mathcal{M}$, and a point $x \in \mathcal{N}$, we denote $\mathcal{C}_x f$ the type (2,0) tensor field over $\mathcal{M}$ given by*

$$\mathcal{C}_x f = (\partial_{x^I} f) \partial_p^2 v^{(I)} .$$

The metrically equivalent type (0,2) tensor field is $(\mathcal{C}_x f)_* = (\partial_{x^I} f) \nabla_p^2 v^{(I)}$.

Next let B, C denote the symmetric bilinear forms on $\mathbb{R}^N \times \mathbb{R}^N$ given by

$$\begin{aligned} B(\xi, \eta) &= B^{IJ} \xi_I \eta_J , \quad B^{IJ} = g(\mathrm{div}_p \, \partial_p^2 v^{(I)}, \mathrm{div}_p \, \partial_p^2 v^{(J)}) , \\ C(\xi, \eta) &= C^{IJ} \xi_I \eta_J , \quad C^{IJ} = \partial_p^2 v^{(I)} \cdot \nabla_p^2 v^{(J)} , \end{aligned}$$

and observe that

$$B^{IJ} \partial_{x^I} h \partial_{x^J} h = g(\mathcal{B}_x h, \mathcal{B}_x h) , \quad C^{IJ} \partial_{x^I} h \partial_{x^J} h = \mathcal{C}_x h \cdot (\mathcal{C}_x h)_* \tag{25}$$

Assumption 4. *We assume that there exist two constants $\beta, \gamma \geq 0$ such that*

$$B(\xi, \xi) \leq \beta A(\xi, \xi) , \qquad C(\xi, \xi) \leq \gamma A(\xi, \xi) , \quad \textit{for all } \xi \in \mathbb{R}^N .$$

The previous assumptions suffice if W lies in the kernel of the Hessian matrix of $v^{(I)}$, i.e., $\nabla_p^2 v^{(I)}(W, \cdot) = 0$, for all $I = 1, \ldots, N$. If this is not the case, we need a last assumption.

Assumption 5. *If $\nabla_p^2 v^{(I)}(W, \cdot) = 0$, for all $I = 1, \ldots, N$ does not hold, then we assume that there exists a constant $\omega > 0$ such that*

$$g(W, W) \leq \omega . \tag{26}$$

Remark. For instance, the bound (26) does not hold for the classical Fokker-Planck equation (8), but $\nabla_p^2 v^{(I)} = 0$ and therefore we need only Assumptions 1–3 in this case.

Remark. For the Fokker-Planck equation (9) we have the following identifications: All indexes (small and capital) run from 1 to $N = M$ and

$$\begin{aligned} & g_{ij} = \delta_{ij} , \quad v^{(I)}(p) = p^I , \quad E = |p|^2/2 ; \\ & W^i = -p^i , \quad \widetilde{\mathrm{Ric}}_{ij} = \delta_{ij} ; \\ & \partial_p^2 v^{(I)} = \mathcal{B}_x f = \mathcal{C}_x f = 0 , \quad \text{for all functions } f ; \\ & A_{IJ} = \delta_{IJ} , \quad Q = W . \end{aligned}$$

The constants in the assumptions can be chosen as $\sigma_1 = \sigma_2 = \alpha = 1, \beta = \gamma = 0$.

Before any claim on the asymptotic time behavior of solutions could be made, one has to ensure that the Cauchy problem for (11) is globally well-posed. In the Appendix we prove that the closure of the operator $A = -L + T$, defined for instance in the space of C^∞ functions with compact support in the p variable, generates a contraction semigroup in $L^2(dxd\mu)$, *provided the dimensions of $\mathcal{N}$ and $\mathcal{M}$ are the same*, i.e. $N = M$, which is the interesting case anyway. In the general case, it will be assumed that the equation (11) admits a unique probability density solution, which is smooth and rapidly decreasing in the p variable, corresponding to a non-negative initial datum with the same regularity and normalized to a probability distribution, i.e.,

$$\|h_{\rm in}\|_{L^1(dxd\mu)} := \int_{\mathbb{T}^N \times \mathbb{R}^M} h_{\rm in}\, dx\, d\mu = 1\,. \tag{27}$$

Our main result is the following.

Theorem 1. *Let the Assumptions 1–5 be verified (Assumptions 1–4 suffice when $\nabla_p^2 v^{(I)}(W,\cdot) \equiv 0$ and Assumptions 1–3 suffice when $\nabla_p^2 v^{(I)} \equiv 0$) and let the initial datum satisfy* (27)*. There exists a constant $C > 0$, depending on suitable norms of $h_{\rm in}$, and a constant $\lambda > 0$, depending on the constants $\sigma_1, \sigma_2, \omega, \alpha, \beta, \gamma$, and which can be explicitly computed, such that the entropy functional*

$$\mathcal{D}[h] = \int_{\mathbb{T}^N \times \mathbb{R}^M} h \log h\, dx\, d\mu$$

satisfies

$$\mathcal{D}[h](t) \le C e^{-\lambda t}\,.$$

Remark. Using (7) we have

$$\|h - 1\|_{L^1(dxd\mu)} = O(e^{-\lambda\, t/2})\,, \qquad \text{as } t \to \infty\,.$$

Equivalently, the solution of (9) converges to the steady state $f_\infty \sim e^{-E}$ with exponential rate in the L^1 norm.

Remark. Of course there is no loss of generality in restricting to initial data that satisfy (27), since (11) preserves the $L^1(dxd\mu)$-norm and is invariant by the rescaling $h \to Mh$. Positive solutions with mass $M > 0$ converge to the equilibrium state $h_\infty = M$.

We complement Theorem 1 with a result that simplifies Assumption 3 in some interesting applications.

Corollary 1. *Let Assumptions 1 and 4-5 (if necessary) hold and assume further that there exists a function $\zeta : \mathcal{M} \to (0,\infty)$ such that*

$$A_{IJ} = \zeta(p)\delta_{IJ}\,, \tag{28}$$

where A_{IJ} is the matrix inverse of $A^{IJ} = g(\partial_p v^{(I)}, \partial_p v^{(J)})$. If there exists a constant $\kappa_1 < \sigma_1$ such that

$$\nabla_p^2 \zeta(X,Y) \le \frac{\zeta}{N}\kappa_1\, g(X,Y)\,, \quad \textit{for all} \;\; X, Y \in \mathfrak{X}(\mathcal{M}) \tag{29}$$

and in addition

$$\kappa_2 = \sup_{p\in\mathcal{M}} \left(\frac{\Delta_p \zeta}{\zeta} + (N-1)\frac{g(\partial_p\zeta, \partial_p\zeta)}{\zeta^2} \right) < 0 \,, \tag{30}$$

then the same conclusion of Theorem 1 holds.

An example that is covered by Corollary 1 is the Fokker-Planck equation (9) with the classical velocity field $v(p) = p$ and an isotropic diffusion matrix, i.e.

$$D_{ij}(p) = \Pi(p)\delta_{ij} \,,$$

where Π is a positive function. In particular $g^{ij} = \Pi(p)\delta^{ij}$. Thus, since $v^{(I)} = p^I$ (all indexes run from 1 to N in this example), we have

$$A^{IJ} = g^{ij}\partial_{p^i} v^{(I)} \partial_{p^j} v^{(J)} = g^{ij}\delta^I_i \delta^J_j = g^{IJ} = \Pi(p)\delta^{IJ} \,,$$

and so (28) holds with $\zeta(p) = 1/\Pi(p)$.

4 Evolution of the modified Entropy

In the rest of the paper the following abbreviations will be used:

$$\int \cdots \, dx\, d\mu = \int_{\mathbb{T}^N\times\mathbb{R}^M} \cdots \, dx\, d\mu$$

and

$$\bar{h} = \log h \,.$$

Moreover the measure $dx\, d\mu$ will be omitted in the proofs.

Recall that

$$\mathcal{D}[h] = \int h\,\bar{h}\, dx\, d\mu$$

and define

$$\mathfrak{I}_{pp}[h] = \int g(\partial_p h, \partial_p \bar{h})\, dx\, d\mu \,,$$

$$\mathfrak{I}_{xp}[h] = \int g(\mathcal{A}_x h, \partial_p \bar{h})\, dx\, d\mu \,,$$

$$\mathfrak{I}_{xx}[h] = \int g(\mathcal{A}_x h, \mathcal{A}_x \bar{h})\, dx\, d\mu \,.$$

Given four constants $a, b, c, k > 0$, we define the modified entropy as

$$\mathcal{E}[h] = k\,\mathfrak{D}[h] + a\,\mathfrak{I}_{pp}[h] + 2b\,\mathfrak{I}_{xp}[h] + c\,\mathfrak{I}_{xx}[h] \,.$$

The purpose of the rest of the present section is to study the time evolution of the modified entropy, by computing the time derivative of $\mathfrak{D}$, $\mathfrak{I}_{xx}, \mathfrak{I}_{xp}$ and $\mathfrak{I}_{pp}$.

Lemma 5. *The following holds:*

$$\frac{d}{dt}\mathfrak{D} = -\mathfrak{I}_{pp} \,,$$

Proof. We compute

$$\frac{d}{dt}\mathfrak{D} = \int \partial_t h(1+\bar{h}) = -\int (1+\bar{h})Th + \int Lh\,dx\,d\mu + \int \bar{h}Lh \,.$$

By (17) and (18), the first two terms vanish and

$$\int \bar{h}Lh = -\int g(\partial_p h, \partial_p \bar{h}) \,.$$

□

Lemma 6. *The following holds:*

$$\begin{aligned}\frac{d}{dt}\mathfrak{I}_{pp} =& -2\mathfrak{I}_{xp} - 2\int h\,\widetilde{\mathrm{Ric}}(\partial_p\bar{h}, \partial_p\bar{h})\,dx\,d\mu \\ & -2\int h\,\partial_p^2\bar{h}\cdot\nabla_p^2\bar{h}\,dx\,d\mu \,.\end{aligned}$$

where $\widetilde{\mathrm{Ric}}$ *is the Bakry-Emery-Ricci tensor* (20)*.*

Proof. We compute

$$\begin{aligned}\frac{d}{dt}\mathfrak{I}_{pp} =& 2\int g(\partial_p\bar{h}, \partial_p\partial_t h) - \int g(\partial_p\bar{h}, \partial_p\bar{h})\partial_t h \\ =& \underbrace{-2\int g(\partial_p\bar{h}, \partial_p(Th))}_{\heartsuit} + \underbrace{\int g(\partial_p\bar{h}, \partial_p\bar{h})Th}_{\diamondsuit} \\ & \underbrace{+2\int g(\partial_p\bar{h}, \partial_p(Lh))}_{\clubsuit} - \underbrace{\int g(\partial_p\bar{h}, \partial_p\bar{h})Lh}_{\spadesuit} \,.\end{aligned}$$

We claim that $\heartsuit + \diamondsuit = -2\mathfrak{I}_{xp}$. We prove this using the coordinates representation. From one hand

$$g(\partial_p\bar{h}, \partial_p(Th)) = g^{ij}\partial_{p^i}\bar{h}(\partial_{p^j}v^{(I)})\partial_{x^I}h + g^{ij}\partial_{p^i}\bar{h}\,v^{(I)}\partial_{p^j}\partial_{x^I}h \,;$$

on the other hand, integrating by parts in the x variable,

$$\diamondsuit = -2\int g^{ij}\partial_{p^j}h\partial_{p^i}\partial_{x^I}\bar{h}\,v^{(I)} = 2\int g^{ij}\partial_{p^j}\partial_{x^I}h\,v^{(I)}\partial_{p^i}\bar{h} \,.$$

Thus

$$\heartsuit + \diamondsuit = -2\int g^{ij}\partial_{p^i}\bar{h}(\partial_{p^j}v^{(I)})\partial_{x^I}h = -2\int g(\mathcal{A}_x h, \partial_p\bar{h}) \,. \qquad (31)$$

The term $\clubsuit$ is

$$\clubsuit = 2\int g(\partial_p\bar{h}, \partial_p(\Delta_p h)) + 2\int g(\partial_p\bar{h}, \partial_p(Wh)) = \clubsuit_1 + \clubsuit_2\,.$$

By (v) of Lemma 1 and (19) we have

$$\begin{aligned}\clubsuit_1 &= -\,2\int \mathrm{Ric}(\partial_p\bar{h}, \partial_p h) + 2\int g(\partial_p\bar{h}, \mathrm{div}_p\,\partial_p^2 h)\\ &= -\,2\int \mathrm{Ric}(\partial_p\bar{h}, \partial_p h) - 2\int \partial_p^2 h\cdot\nabla_p^2\bar{h} - 2\int \partial_p^2 h(W_*, \nabla_p\bar{h})\,.\end{aligned}$$

Moreover by (iv) of Lemma 1,

$$\clubsuit_2 = 2\int g(\partial_p\bar{h}, \partial_p(g(\partial_p h, W))) = 2\int \partial_p^2 h(W_*, \nabla_p\bar{h}) + 2\int \partial_p W(\nabla_p h, \nabla_p\bar{h})\,.$$

Summing up and using the identity

$$\partial_p^2 h = h\,\partial_p^2\bar{h} + \partial_p\bar{h}\otimes\partial_p h$$

we obtain

$$\clubsuit = -2\int h\,\widetilde{\mathrm{Ric}}(\partial_p\bar{h}, \partial_p\bar{h}) - 2\int h\,\partial_p^2\bar{h}\cdot\nabla_p^2\bar{h} - 2\int \nabla_p^2\bar{h}(\partial_p\bar{h}, \partial_p h)\,. \tag{32}$$

Finally, by (18) and (iv) of Lemma 1,

$$\spadesuit = \int g(\partial_p h, \partial_p(g(\partial_p\bar{h}, \partial_p\bar{h}))) = 2\int \partial_p^2\bar{h}(\nabla_p\bar{h}, \nabla_p h)\,, \tag{33}$$

which cancels out with the last term of (32). The claim follows summing up (31)–(33). $\square$

Lemma 7. *The following holds:*

$$\begin{aligned}\frac{d}{dt}\mathfrak{I}_{xp} = &-\mathfrak{I}_{xx} - \int h\,\widetilde{\mathrm{Ric}}(\mathcal{A}_x\bar{h}, \partial_p\bar{h})\,dx\,d\mu\\ &- 2\int h\,\partial_p^2\bar{h}\cdot\nabla_p(\mathcal{A}_x\bar{h})_*\,dx\,d\mu + \int g(\partial_p\bar{h}, \mathcal{B}_x h)\,dx\,d\mu\\ &+ 2\int h\,\partial_p^2\bar{h}\cdot(\mathcal{C}_x\bar{h})_*\,dx\,d\mu + \int \partial_p^2 v^{(I)}(W_*, \partial_{x^I} h\nabla_p\bar{h})\,dx\,d\mu\,.\end{aligned}$$

Proof. We have

$$\begin{aligned}\frac{d}{dt}\mathfrak{I}_{xp} &= \int g(\mathcal{A}_x\partial_t h, \partial_p\bar{h}) - \int g(\mathcal{A}_x\bar{h}, \partial_p\bar{h})\partial_t h + \int g(\mathcal{A}_x\bar{h}, \partial_p\partial_t h))\\ &= \underbrace{-\int g(\mathcal{A}_x\bar{h}, \partial_p(Th))}_{\heartsuit_1}\underbrace{-\int g(\mathcal{A}_x(Th), \partial_p\bar{h})}_{\heartsuit_2}\underbrace{+\int g(\mathcal{A}_x\bar{h}, \partial_p\bar{h})Th}_{\heartsuit_3}\\ &\quad\underbrace{-\int g(\mathcal{A}_x\bar{h}, \partial_p\bar{h})Lh}_{\diamondsuit}\underbrace{+\int g(\mathcal{A}_x\bar{h}, \partial_p(Lh))}_{\clubsuit}\underbrace{+\int g(\mathcal{A}_x(Lh), \partial_p\bar{h})}_{\spadesuit}\,.\end{aligned}$$

Now we claim that

$$\heartsuit = \heartsuit_1 + \heartsuit_2 + \heartsuit_3 = -\mathfrak{I}_{xx} \; . \tag{34}$$

In fact, using the coordinates representation the first term of $\heartsuit$ can be rewritten as

$$\begin{aligned}\heartsuit_1 &= -\int g^{ij} \partial_{p^i} v^{(I)} \partial_{x^I} \bar{h}\, \partial_{p^j} v^{(J)} \partial_{x^J} h - \int g^{ij} \partial_{p^i} v^{(I)} \partial_{x^I} \bar{h}\, v^{(J)} \partial_{p^j} \partial_{x^J} h \\ &= \heartsuit_{1A} + \heartsuit_{1B} \; .\end{aligned}$$

It is clear that $\heartsuit_{1A} = -\mathfrak{I}_{xx}$. Integrating by parts in the x variable we obtain

$$\heartsuit_{1B} = \int g^{ij} \partial_{p^i} v^{(I)} \partial_{x^I} \partial_{x^J} \bar{h}\, v^{(J)} \partial_{p^j} h \; .$$

In the previous expression we use the identity

$$\partial_{x^I} \partial_{x^J} \bar{h} = h^{-1} \partial_{x^I} \partial_{x^J} h - h^{-2} \partial_{x^I} h \partial_{x^J} h$$

and so doing we obtain

$$\begin{aligned}\heartsuit_{1B} &= \int g^{ij} \partial_{p^i} v^{(I)} \partial_{x^I} \partial_{x^J} h\, v^{(J)} \partial_{p^j} \bar{h} - \int g^{ij} \partial_{p^i} v^{(I)} \partial_{x^I} \bar{h}\, v^{(J)} \partial_{x^J} h \partial_{p^j} \bar{h} \\ &= \int g(\mathcal{A}_x(Th), \partial_p \bar{h}) - \int g(\mathcal{A}_x \bar{h}, \partial_p \bar{h}) Th = -\heartsuit_2 - \heartsuit_3 \; .\end{aligned}$$

This proves (34). It remains to study the integrals $\diamondsuit, \clubsuit, \spadesuit$. We begin by applying (18) and (iv) of Lemma 1 to $\diamondsuit$:

$$\diamondsuit = \int g(\partial_p h, \partial_p(g(\mathcal{A}_x \bar{h}, \partial_p \bar{h}))) = \int \partial_p(\mathcal{A}_x \bar{h})(\nabla_p \bar{h}, \nabla_p h) + \int \partial_p^2 \bar{h}((\mathcal{A}_x \bar{h})_*, \nabla_p h) \; . \tag{35}$$

As to $\clubsuit$, we first split it as

$$\clubsuit = \int g(\mathcal{A}_x \bar{h}, \partial_p(\Delta_p h)) + \int g(\mathcal{A}_x \bar{h}, \partial_p(Wh)) = \clubsuit_1 + \clubsuit_2 \; .$$

By (v) of Lemma 1 and (19) we have

$$\begin{aligned}\clubsuit_1 &= \int g(\mathcal{A}_x \bar{h}, \mathrm{div}_p\, \partial_p^2 h) - \int \mathrm{Ric}(\mathcal{A}_x \bar{h}, \partial_p h) \\ &= -\int \partial_p^2 h \cdot \nabla_p (\mathcal{A}_x \bar{h})_* - \int \partial_p^2 h(W_*, (\mathcal{A}_x \bar{h})_*) - \int \mathrm{Ric}(\mathcal{A}_x \bar{h}, \partial_p h) \; .\end{aligned}$$

Likewise

$$\clubsuit_2 = \int g(\mathcal{A}_x \bar{h}, \partial_p(g(\partial_p h, W))) = \int \partial_p^2 h(W_*, (\mathcal{A}_x \bar{h})_*) + \int \nabla_p W_*(\partial_p h, \mathcal{A}_x \bar{h}) \; .$$

Summing up,

$$\clubsuit = -\int \widetilde{\mathrm{Ric}}(\mathcal{A}_x \bar{h}, \partial_p h) - \int \partial_p^2 h \cdot \nabla_p (\mathcal{A}_x \bar{h})_* \; .$$

In the second integral we replace

$$\partial_p^2 h = h\,\partial_p^2\bar{h} + \partial_p\bar{h}\otimes\partial_p h$$

and we get

$$♣ = -\int h\,\widetilde{\mathrm{Ric}}(\mathcal{A}_x\bar{h},\partial_p\bar{h}) - \int h\,\partial_p^2\bar{h}\cdot\nabla_p(\mathcal{A}_x\bar{h})_* - \int \partial_p(\mathcal{A}_x\bar{h})(\nabla_p\bar{h},\nabla_p h)\,. \quad (36)$$

Note the the last term in the r.h.s. of (36) cancels out with the first term in the r.h.s. of (35). We now work out the term ♠. First by means of an integration by parts in the x variable we can rewrite it as

$$♠ = -\int (Lh)g(\partial_p v^{(I)},\partial_p(\partial_{x^I}\bar{h}))\,.$$

Then by (18) and (iv) of Lemma 1 we have

$$♠ = \int \partial_p^2 v^{(I)}(\nabla_p\partial_{x^I}\bar{h},\nabla_p h) + \int \partial_p^2(\partial_{x^I}\bar{h})(\nabla_p v^{(I)},\nabla_p h) = ♠_1 + ♠_2\,. \quad (37a)$$

In $♠_1$ we integrate by parts in x, apply the identity

$$\nabla_p\bar{h}\otimes\nabla_p\partial_{x^I}h = \nabla_p(\nabla_p\bar{h}\partial_{x^I}h) - \nabla_p^2\bar{h}\partial_{x^I}h$$

and (19) to obtain

$$\begin{aligned}
♠_1 &= -\int \partial_p^2 v^{(I)}(\nabla_p\bar{h},\nabla_p\partial_{x^I}h) = -\int \partial_p^2 v^{(I)}\cdot\nabla_p\bar{h}\otimes\nabla_p\partial_{x^I}h\\
&= \int \partial_p^2 v^{(I)}\partial_{x^I}h\cdot\nabla_p^2\bar{h} - \int \partial_p^2 v^{(I)}\cdot\nabla_p(\nabla_p\bar{h}\partial_{x^I}h)\\
&= \int h\,\partial_p^2\bar{h}\cdot(\mathcal{C}_x\bar{h})_* + \int g(\partial_p\bar{h},\mathcal{B}_x h) + \int \partial_p^2 v^{(I)}(W_*,\partial_{x^I}h\nabla_p\bar{h})\,. \quad (37b)
\end{aligned}$$

In $♠_2$ we integrate by parts in the x variable and apply the identity

$$\nabla_p\partial_{x^I}h = h\nabla_p\partial_{x^I}\bar{h} + \partial_{x^I}\bar{h}\nabla_p h$$

to obtain

$$\begin{aligned}
♠_2 &= -\int \partial_p^2\bar{h}(\nabla_p v^{(I)},\nabla_p\partial_{x^I}h)\\
&= -\int \partial_p^2\bar{h}(\nabla_p v^{(I)},\partial_{x^I}\bar{h}\nabla_p h) - \int h\,\partial_p^2\bar{h}(\nabla_p v^{(I)},\nabla_p\partial_{x^I}\bar{h})\\
&= -\int \partial_p^2\bar{h}((\mathcal{A}_x\bar{h})_*,\nabla_p h) - \int h\,\partial_p^2\bar{h}\cdot\nabla_p v^{(I)}\otimes\nabla_p\partial_{x^I}\bar{h} = ♠_{2A} + ♠_{2B}\,.
\end{aligned} \quad (37c)$$

Note that $♠_{2A}$ cancels out with the second term in the r.h.s. of (35). In $♠_{2B}$ we use

$$\nabla_p v^{(I)}\otimes\nabla_p\partial_{x^I}\bar{h} = \nabla_p(\mathcal{A}_x\bar{h})_* - \nabla_p^2 v^{(I)}\partial_{x^I}\bar{h}$$

to finally obtain

$$\spadesuit_{2B} = -\int h\,\partial_p^2\bar{h}\cdot\nabla_p(\mathcal{A}_x\bar{h})_* + \int h\,\partial_p^2\bar{h}\cdot(\mathcal{C}_x\bar{h})_*\,. \tag{37d}$$

The claim follows by (34)–(37). □

Lemma 8. *The following holds:*

$$\begin{aligned}\frac{d}{dt}\mathfrak{I}_{xx} =& -2\int h\,\partial_p(\mathcal{A}_x\bar{h})\cdot\nabla_p(\mathcal{A}_x\bar{h})_*\,dx\,d\mu + 2\int g(\mathcal{A}_x\bar{h},\mathcal{B}_x h)\,dx\,d\mu \\ &+ 4\int h\,\partial_p(\mathcal{A}_x\bar{h})\cdot(\mathcal{C}_x h)_*\,dx\,d\mu + 2\int \partial_p^2 v^{(I)}(W_*,\partial_{x^I}h\,(\mathcal{A}_x\bar{h})_*)\,dx\,d\mu\,.\end{aligned}$$

Proof. The proof is very similar to that of Lemma 7. First we compute

$$\begin{aligned}\frac{d}{dt}\mathfrak{I}_{xx} =& 2\int g(\mathcal{A}_x\partial_t h,\mathcal{A}_x\bar{h}) - \int g(\mathcal{A}_x\bar{h},\mathcal{A}_x\bar{h})\partial_t h \\ =& \underbrace{-2\int g(\mathcal{A}_x(Th),\mathcal{A}_x\bar{h})}_{\heartsuit} + \underbrace{\int Th\,g(\mathcal{A}_x\bar{h},\mathcal{A}_x\bar{h})}_{\diamondsuit} \\ & \underbrace{-\int Lh\,g(\mathcal{A}_x\bar{h},\mathcal{A}_x\bar{h})}_{\clubsuit} + \underbrace{2\int g(\mathcal{A}_x(Lh),\mathcal{A}_x\bar{h})}_{\spadesuit}\end{aligned}$$

We claim that

$$\heartsuit + \diamondsuit = 0\,. \tag{38}$$

In fact, by (17)

$$\begin{aligned}\diamondsuit &= -\int hT(g(\mathcal{A}_x\bar{h},\mathcal{A}_x\bar{h})) = -2\int hg(\mathcal{A}_x(T\bar{h}),\mathcal{A}_x\bar{h}) \\ &= -2\int g(\mathcal{A}_x(Th),\mathcal{A}_x\bar{h}) + 2\int (Th)\,g(\mathcal{A}_x\bar{h},\mathcal{A}_x\bar{h}) \Rightarrow \diamondsuit = -\heartsuit\,.\end{aligned}$$

By (18) and (iv) of Lemma 1 the term ♣ can be rewritten as

$$\clubsuit = 2\int \partial_p(\mathcal{A}_x\bar{h})((\mathcal{A}_x\bar{h})_*,\nabla_p h)\,. \tag{39}$$

Likewise,

$$\begin{aligned}\spadesuit &= -2\int Lh\,g(\partial_p v^{(I)},\mathcal{A}_x(\partial_{x^I}\bar{h})) = 2\int g(\partial_p h,\partial_p(g(\partial_p v^{(I)},\mathcal{A}_x(\partial_{x^I}\bar{h})))) \\ &= 2\int \partial_p^2 v^{(I)}(\mathcal{A}_x(\partial_{x^I}\bar{h})_*,\nabla_p h) + 2\int \partial_p(\mathcal{A}_x(\partial_{x^I}\bar{h}))(\nabla_p v^{(I)},\nabla_p h) \\ &= \spadesuit_1 + \spadesuit_2\,. \end{aligned} \tag{40a}$$

Integrating by parts in x and using the identity

$$(\mathcal{A}_x\bar{h})_* \otimes \nabla_p \partial_{x^I} h = \nabla_p((\mathcal{A}_x\bar{h})_* \partial_{x^I} h) - \partial_{x^I} h \nabla_p (\mathcal{A}_x\bar{h})_* ,$$

we may rewrite $\spadesuit_1$ as

$$\begin{aligned}\spadesuit_1 &= -2\int \partial_p^2 v^{(I)}((\mathcal{A}_x\bar{h})_*, \nabla_p\partial_{x^I} h) = -2\int \partial_p^2 v^{(I)} \cdot (\mathcal{A}_x\bar{h})_* \otimes \nabla_p \partial_{x^I} h \\ &= -2\int \partial_p^2 v^{(I)} \cdot \nabla_p((\mathcal{A}_x\bar{h})_* \partial_{x^I} h) + 2\int \partial_p^2 v^{(I)} \cdot \nabla_p(\mathcal{A}_x\bar{h})_* \partial_{x^I} h .\end{aligned}$$

Applying (19) to the first term in the last line we get

$$\spadesuit_1 = 2\int g(\mathcal{A}_x\bar{h}, \mathcal{B}_x h) + 2\int \partial_p^2 v^{(I)}(W_*, (\mathcal{A}_x\bar{h})_* \partial_{x^I} h) + 2\int h\, \partial_p(\mathcal{A}_x\bar{h}) \cdot (\mathcal{C}_x\bar{h})_* . \tag{40b}$$

Integrating by parts in the x variable and using the identity

$$\nabla_p \partial_{x^I} h = h \nabla_p \partial_{x^I} \bar{h} + \partial_{x^I} \bar{h} \nabla_p h$$

the term $\spadesuit_2$ becomes

$$\begin{aligned}\spadesuit_2 &= -2\int h\, \partial_p(\mathcal{A}_x\bar{h})(\nabla_p v^{(I)}, \nabla_p \partial_{x^I}\bar{h}) - 2\int \partial_p(\mathcal{A}_x\bar{h})(\nabla_p v^{(I)}, \partial_{x^I}\bar{h}\nabla_p h) \\ &= \spadesuit_{2A} + \spadesuit_{2B} .\end{aligned} \tag{40c}$$

Note that $\spadesuit_{2B}$ cancels out with $\clubsuit$. In $\spadesuit_{2A}$ we use

$$\partial_p(\mathcal{A}_x\bar{h})(\nabla_p v^{(I)}, \nabla_p \partial_{x^I}\bar{h}) = \partial_p(\mathcal{A}_x\bar{h}) \cdot \nabla_p v^{(I)} \otimes \nabla_p \partial_{x^I}\bar{h}$$

and

$$\nabla_p v^{(I)} \otimes \nabla_p \partial_{x^I}\bar{h} = \nabla_p(\mathcal{A}_x\bar{h})_* - \nabla_p^2 v^{(I)} \partial_{x^I}\bar{h}$$

to obtain

$$\spadesuit_{2A} = -2\int h\, \partial_p(\mathcal{A}_x\bar{h}) \cdot \nabla_p(\mathcal{A}_x\bar{h}_*) + 2\int h\, \partial_p(\mathcal{A}_x\bar{h}) \cdot (\mathcal{C}_x h)_* . \tag{40d}$$

Summing up (38)–(40) concludes the proof. □

Remark. If $\nabla_p^2 v^{(I)}(W, \cdot) = 0$, cf. Assumption 5, the last terms in $d\mathfrak{I}_{xp}/dt$ and $d\mathfrak{I}_{xx}/dt$ vanish. If this is not the case, we use that

$$\partial_p^2 v^{(I)}(W_*, \partial_{x^I} h \nabla_p \bar{h}) = \mathcal{C}_x h(W_*, \nabla_p \bar{h})$$

in Lemma 7 and

$$\partial_p^2 v^{(I)}(W_*, \partial_{x^I} h (\mathcal{A}_x\bar{h})_*) = \mathcal{C}_x h(W_*, (\mathcal{A}_x\bar{h})_*)$$

in Lemma 8. We shall continue the proof assuming that these two terms do not vanish, which is the most difficult case anyway.

5 A differential inequality for the modified entropy

Recall that

$$\mathcal{E}[h] = k\,\mathfrak{D}[h] + a\,\mathfrak{I}_{pp}[h] + 2b\,\mathfrak{I}_{xp}[h] + c\,\mathfrak{I}_{xx}[h]\ .$$

In this section we prove that, under suitable conditions on the constants a, b, c, k, the modified entropy satisfies

$$\mathcal{E}[h] \geq k\,\mathfrak{D}[h]\ , \qquad \frac{d}{dt}\mathcal{E}[h] \leq -d(a\,\mathfrak{I}_{xx} + 2b\,\mathfrak{I}_{xp} + c\,\mathfrak{I}_{pp})\ , \tag{41}$$

where d is a positive constant. In particular, the first bound shows that exponential decay of the modified entropy implies exponential decay of the entropy.

The bound from below is easily established.

Lemma 9. *Assume $b \leq \sqrt{ac}$. Then $\mathcal{E}[h] \geq k\,\mathfrak{D}[h]$.*

Proof. By Young's inequality, for all $\varepsilon > 0$ we have

$$g(\mathcal{A}_x h, \partial_p \bar{h}) \geq -\varepsilon g(\partial_p h, \partial_p \bar{h}) - \frac{1}{4\varepsilon} g(\mathcal{A}_x h, \mathcal{A}_x \bar{h})\ ,$$

whence $2b\mathfrak{I}_{xp} \geq -2b\varepsilon\mathfrak{I}_{pp} - (b/2\varepsilon)\mathfrak{I}_{xx}$ and so

$$\mathcal{E}[h] \geq k\,\mathfrak{D}[h] + \left(c - \frac{b}{2\varepsilon}\right)\mathfrak{I}_{xx}[h] + (a - 2\varepsilon b)\mathfrak{I}_{pp}[h] \geq k\,\mathfrak{D}[h]\ ,$$

provided $b/c \leq 2\varepsilon \leq a/b$. □

The bound from above, which requires the assumptions of the main theorem (except Assumption 3), is more complicated. Since

$$a\,\mathfrak{I}_{xx} + 2b\,\mathfrak{I}_{xp} + c\,\mathfrak{I}_{pp} \leq \max(a+b, b+c)(\mathfrak{I}_{xx} + \mathfrak{I}_{pp})\ ,$$

it suffices to prove the following.

Proposition 1. *Let the Assumptions 1, 2, 4, 5 hold. There exists a region $\Omega \subset \mathbb{R}^3$ such that $\Omega \subset \{(a,b,c) : b \leq \sqrt{ac}\}$ and, for all $(a,b,c) \in \Omega$, there exists $d > 0$ such that*

$$\frac{d}{dt}\mathcal{E}[h] \leq -d\big(\mathfrak{I}_{xx}[h] + \mathfrak{I}_{pp}[h]\big)\ . \tag{42}$$

Remark. The best constant in the inequality (42) may be written as $\bar{d} = \sup_{\bar{\Omega}} d$, where $\bar{\Omega}$ is the largest region for which Proposition 1 holds. We refrain from computing it explicitly, since the method we use is anyway unsuitable to obtain the optimal rate of decay of the entropy.

The proof of the proposition is based on the following lemma.

Lemma 10. *For all constants $\varepsilon_1, \dots, \varepsilon_{10} > 0$ we have*

$$\frac{d}{dt}\mathfrak{I}_{pp} \le 2\varepsilon_1 \mathfrak{I}_{xx} + \left(\frac{1}{2\varepsilon_1} - 2\sigma_1\right)\mathfrak{I}_{pp} - 2Q_{pp}^2\,, \tag{43}$$

$$\begin{aligned}\frac{d}{dt}\mathfrak{I}_{xp} \le &\Big[\varepsilon_2\sigma + \varepsilon_3\sigma_1 + (2\varepsilon_5 + \varepsilon_7)\gamma + \varepsilon_6\beta - 1\Big]\mathfrak{I}_{xx} \\ &+ \frac{1}{4}\left(\frac{\sigma}{\varepsilon_2} + \frac{\sigma_1}{\varepsilon_3} + \frac{1}{\varepsilon_6} + \frac{\omega}{\varepsilon_7}\right)\mathfrak{I}_{pp} + \left(2\varepsilon_4 + \frac{1}{2\varepsilon_5}\right)Q_{pp}^2 + \frac{1}{2\varepsilon_4}Q_{xp}^2\,,\end{aligned} \tag{44}$$

$$\frac{d}{dt}\mathfrak{I}_{xx} \le \left(4\varepsilon_8\gamma + \frac{1}{2\varepsilon_9} + 2\varepsilon_9\beta + 2\varepsilon_{10}\gamma + \frac{\omega}{2\varepsilon_{10}}\right)\mathfrak{I}_{xx} + \left(\frac{1}{\varepsilon_8} - 2\right)Q_{xp}^2\,, \tag{45}$$

where

$$Q_{pp}^2 = \int h\,\partial_p^2\bar{h}\cdot\nabla_p^2\bar{h}\,, \quad Q_{xp}^2 = \int h\,\partial_p(\mathcal{A}_x\bar{h})\cdot\nabla_p(\mathcal{A}_x\bar{h}_*)\,.$$

Proof. The inequality (43) is a straightforward consequence of Lemma 6, the inequality $\mathfrak{I}_{xp} \ge -\varepsilon_1\mathfrak{I}_{xx} - (4\varepsilon_1)^{-1}\mathfrak{I}_{pp}$, and Assumption 1. We now prove (44). Using the identity

$$\begin{aligned}\widetilde{\mathrm{Ric}}(\mathcal{A}_x\bar{h}, \partial_p\bar{h}) =&\widetilde{\mathrm{Ric}}\left(\sqrt{\varepsilon_2}\mathcal{A}_x\bar{h} + \frac{1}{\sqrt{4\varepsilon_2}}\partial_p\bar{h}, \sqrt{\varepsilon_2}\mathcal{A}_x\bar{h} + \frac{1}{\sqrt{4\varepsilon_2}}\partial_p\bar{h},\right) \\ &- \varepsilon_2\widetilde{\mathrm{Ric}}(\mathcal{A}_x\bar{h}, \mathcal{A}_x\bar{h}) - \frac{1}{4\varepsilon_2}\widetilde{\mathrm{Ric}}(\partial_p\bar{h}, \partial_p\bar{h})\,,\end{aligned}$$

together with Assumption 1 and $\mathfrak{I}_{xp} \ge -\varepsilon_3\mathfrak{I}_{xx} - (4\varepsilon_3)^{-1}\mathfrak{I}_{pp}$, we get

$$-\int h\,\widetilde{\mathrm{Ric}}(\mathcal{A}_x\bar{h}, \partial_p\bar{h}) \le (\varepsilon_2\sigma + \varepsilon_3\sigma_1)\mathfrak{I}_{xx} + \left(\frac{\sigma}{4\varepsilon_2} + \frac{\sigma_1}{4\varepsilon_3}\right)\mathfrak{I}_{pp}\,. \tag{46}$$

By Young's inequality

$$-2\int h\,\partial_p^2\bar{h}\cdot\nabla_p(\mathcal{A}_x\bar{h})_* \le 2\varepsilon_4\int h\,\partial_p^2\bar{h}\cdot\nabla_p^2\bar{h} + \frac{1}{2\varepsilon_4}\int h\,\partial_p(\mathcal{A}_x\bar{h})\cdot\nabla_p(\mathcal{A}_x\bar{h})_*. \tag{47}$$

By Young's inequality, (25), Assumption 4 and (22)

$$\begin{aligned}2\int h\,\partial_p^2\bar{h}\cdot(\mathcal{C}_x\bar{h})_* &\le 2\varepsilon_5\int h\,\mathcal{C}_x\bar{h}\cdot(\mathcal{C}_x\bar{h})_* + \frac{1}{2\varepsilon_5}\int h\,\partial_p^2\bar{h}\cdot\nabla_p^2\bar{h} \\ &\le 2\varepsilon_5\gamma\mathfrak{I}_{xx} + \frac{1}{2\varepsilon_5}\int h\,\partial_p^2\bar{h}\cdot\nabla_p^2\bar{h}\,.\end{aligned} \tag{48}$$

Likewise

$$\int g(\partial_p\bar{h}, \mathcal{B}_x h) \le \varepsilon_6\int h\,g(\mathcal{B}_x\bar{h}, \mathcal{B}_x\bar{h}) + \frac{1}{4\varepsilon_6}\mathfrak{I}_{pp} \le \varepsilon_6\beta\mathfrak{I}_{xx} + \frac{1}{4\varepsilon_6}\mathfrak{I}_{pp}\,. \tag{49}$$

Finally by (26) and the remark at the end of Section 4,

$$\begin{aligned}\int \partial_p^2 v^{(I)}(W_*, \partial_{x^I} h \nabla_p \bar{h}) &= \int \mathcal{C}_x h(W_*, \nabla_p \bar{h}) = \int h\, \mathcal{C}_x \bar{h} \cdot W_* \otimes \nabla_p \bar{h} \\ &\le \varepsilon_7 \int h\, \mathcal{C}_x \bar{h} \cdot (\mathcal{C}_x \bar{h})_* + \frac{1}{4\varepsilon_7} \int h\, W \otimes \partial_p \bar{h} \cdot W_* \otimes \nabla_p \bar{h} \\ &= \varepsilon_7 \int h\, \mathcal{C}_x \bar{h} \cdot (\mathcal{C}_x \bar{h})_* + \frac{1}{4\varepsilon_7} \int g(W,W) g(\partial_p h, \partial_p \bar{h}) \\ &\le \varepsilon_7 \gamma \mathfrak{I}_{xx} + \frac{\omega}{4\varepsilon_7} \mathfrak{I}_{pp}\,. \qquad (50)\end{aligned}$$

Using the inequalities (46)–(50) in Lemma 7 concludes the proof of (44). The proof of (45) is similar. Reasoning as before one can prove that

$$\begin{aligned}&4 \int h\, \partial_p(\mathcal{A}_x \bar{h}) \cdot (\mathcal{C}_x h)_* \le 4\varepsilon_8 \gamma \mathfrak{I}_{xx} + \frac{1}{\varepsilon_8} Q_{xp}^2\,, \\ &2 \int g(\mathcal{A}_x \bar{h}, \mathcal{B}_x h) \le \Big(2\varepsilon_9 \beta + \frac{1}{2\varepsilon_9}\Big) \mathfrak{I}_{xx}\,, \\ &2 \int \partial_p^2 v^{(I)}(W_*, \partial_{x^I} h (\mathcal{A}_x \bar{h})_*) = 2 \int h\, \mathcal{C}_x \bar{h} \cdot W_* \otimes (\mathcal{A}_x \bar{h})_* \le \Big(2\varepsilon_{10}\gamma + \frac{\omega}{2\varepsilon_{10}}\Big) \mathfrak{I}_{xx}\end{aligned}$$

and substituting in Lemma 8 completes the proof. □

Remark. We are going to apply Lemma 10 for special values of the constants $\varepsilon_1, \dots, \varepsilon_{10}$. In its generality, Lemma 10 could be useful to improve the constant d in (42).

Proof of Proposition 1. In the inequalities (43)–(45) we set

$$\begin{aligned}&\varepsilon_1 = (2a)^{-1}\,, \quad \varepsilon_2 = \varepsilon_3 = \varepsilon_6 = \varepsilon_7 = \frac{1}{4}(\sigma_2 + \beta + \gamma)^{-1}\,, \\ &\varepsilon_4 = \frac{8}{7}(1 + \beta + 17\gamma + \omega)\,, \quad \varepsilon_5 = (8\gamma)^{-1}\,, \quad \varepsilon_8 = 4\,,\ \varepsilon_9 = \varepsilon_{10} = \frac{1}{2}\,.\end{aligned}$$

So doing we obtain

$$\begin{aligned}&\dot{\mathfrak{I}}_{pp} \le a^{-1} \mathfrak{I}_{xx} + (a - 2\sigma_1) \mathfrak{I}_{pp} - 2Q_{pp}^2\,, \\ &\dot{\mathfrak{I}}_{xp} \le -\frac{1}{2} \mathfrak{I}_{xx} + s_1 s_2 \mathfrak{I}_{pp} + \left(\frac{16s}{7} + 4\gamma\right) Q_{pp}^2 + \frac{7}{16s} Q_{xp}^2\,, \\ &\dot{\mathfrak{I}}_{xx} \le s \mathfrak{I}_{xx} - \frac{7}{4} Q_{xp}^2\,,\end{aligned}$$

where

$$s_1 = \sigma_2 + \beta + \gamma\,,\ s_2 = 1 + \sigma_2 + \omega\,,\ s = 1 + \beta + 17\gamma + \omega\,.$$

Therefore

$$\begin{aligned}\frac{d}{dt}\mathcal{E}[h] &= k\,\frac{d}{dt}\mathfrak{D}[h] + a\,\frac{d}{dt}\mathfrak{I}_{pp} + 2b\,\frac{d}{dt}\mathfrak{I}_{xp} + c\,\frac{d}{dt}\mathfrak{I}_{xx} \\ &\le \left[-k + a(a-2\sigma_1) + 2bs_1s_2\right]\mathfrak{I}_{pp} + (1+cs-b)\,\mathfrak{I}_{xx} \\ &\quad + 2\left[b\left(\frac{16}{7}s + 4\gamma\right) - a\right]Q_{pp}^2 + \frac{7}{4}\left(\frac{b}{2s} - c\right)Q_{xp}^2\,.\end{aligned}$$

It is clear that the coefficient of $\mathfrak{I}_{pp}$ can be made negative by choosing k sufficiently large, for all values of the other constants. To make the coefficients of $\mathfrak{I}_{xx}$, Q_{pp}^2, Q_{xp}^2 negative we require that

$$b > 1 + cs\,, \quad b < \frac{a}{\frac{16}{7}s + 4\gamma}\,, \quad b < 2cs\,.$$

This is possible as soon as

$$a > (1+cs)(\frac{16}{7}s + 4\gamma) \qquad \text{and} \qquad c > s^{-1}\,.$$

If we further require that $a > 4s^2c$, then $2cs < \sqrt{ac}$ and therefore $b < 2cs$ implies $b < \sqrt{ac}$ as well. This completes the proof of the proposition. □

6 The Log-Sobolev inequality and decay of the entropy

We shall now prove that the following logarithmic Sobolev inequality holds:

$$\mathfrak{D}[h] \le \frac{1}{2\alpha}\big(\mathfrak{I}_{xx}[h] + \mathfrak{I}_{pp}[h]\big), \tag{51}$$

for all smooth probability densities h, not necessarily solutions of (11), where α is the constant in Assumption 3. Replacing in (42) we obtain

$$\frac{d}{dt}\mathcal{E}[h] \le -(2d\alpha)\mathfrak{D}[h]$$

and combining with the second inequality in (41) we infer that there exists a constant $\lambda > 0$ such that

$$\frac{d}{dt}\mathcal{E}[h] \le -\lambda\mathcal{E}[h]\,.$$

Whence $\mathcal{E}[h] = O(e^{-\lambda t})$ and by the lower bound $\mathcal{E}[h] \ge k\,\mathfrak{D}[h]$, see Lemma 9, the entropy decays exponentially as well, which is the main claim of Theorem 1.

Proposition 2. *Let Assumption 3 be satisfied. Then* (51) *holds.*

Proof. Consider the non-degenerate Fokker-Planck equation

$$\partial_t f = \Delta^G f + Qf \tag{52}$$

on $\mathcal{M} \times \mathcal{N}$. The entropy functional and entropy dissipation functional associated to (52) are exactly $\mathfrak{D}$ and $\mathfrak{I} = \mathfrak{I}_{pp} + \mathfrak{I}_{xx}$. Assumption 3 asserts that the metric G and the vector field Q verify the curvature bound condition. Thus the logarithmic Sobolev inequality (51) holds, as proved in [2]. □

The proof of Theorem 1 is complete. In the rest of this section we prove Corollary 1. Clearly (28) entails that Assumption 2 holds. We shall now prove that under the bounds (29)-(30), Assumption 3 is satisfied as well. To see this, we observe that when A_{IJ} has the form (28), the Riemannian manifold $(\mathcal{M}\times\mathcal{N}, G)$ is the warped product of the manifolds $(\mathbb{R}^M, g)$ and $(\mathbb{T}^N, \delta)$, where δ is the flat Euclidean metric on the torus. See [12, Ch. 14] for an introduction to the geometry of warped product manifolds. In particular, by Corollary 43 of [12, Ch. 14] we have that, for all horizontal (i.e., tangent to $\mathcal{M}$) vector fields X, Y and vertical (i.e. tangent to $\mathcal{N}$) vector fields V, W, the following identities hold:

$$\begin{aligned}
&\mathrm{Ric}^G(X,Y) = \mathrm{Ric}(X,Y) - \frac{N}{\zeta}\nabla_p^2\zeta(X,Y)\,,\\
&\mathrm{Ric}^G(X,V) = 0\,,\\
&\mathrm{Ric}^G(V,W) = -\left(\frac{\Delta_p\zeta}{\zeta} + (N-1)\frac{g(\partial_p\zeta,\partial_p\zeta)}{\zeta^2}\right)G(V,W)\,.
\end{aligned}$$

Thus from the condition (30) we have

$$\mathrm{Ric}^G(V,W) \geq |\kappa_2| G(V,W)\,.$$

Moreover, by Assumption 1 and (29), we also have

$$\mathrm{Ric}^G(X,Y) \geq (\sigma_1 - \kappa_1) G(X,Y)\,.$$

The conclusion of Corollary 1 follows.

A Appendix: Global regularity of solutions

In this appendix we discuss the global regularity of solutions to the equation (11) in the case when the dimensions of the spaces $\mathcal{N}$ and $\mathcal{M}$ coincide, i.e., $N = M$. The argument below generalizes immediately to the case $N \leq M$; when $N > M$ the problem is a bit more tricky and will not be considered here. The following discussion is based on [8, Ch. 5], except that we work in a different functions space. To adhere with the conventions used in [8], we rewrite (11) as

$$\partial_t h + Ah = 0\,,$$

where

$$A = -\Delta_p - W + v(p) \cdot \nabla_x = -L + T \, .$$

The domain $D(A)$ of the operator A is chosen as the space of C^∞ functions on $\mathbb{T}^N \times \mathbb{R}^N$ with compact support in the $p \in \mathbb{R}^N$ variable, which is dense in $\mathcal{H} := L^2(dxd\mu)$. Our purpose is to prove that the closure of A generates a C^∞-regularizing contraction semigroup in $\mathcal{H}$. To this end we need to assume that the quantities g, v, E are C^∞, but using a simple iteration scheme and the argument below, one can prove existence of global regular solutions under milder regularity conditions on g, v, E , such as those imposed in the main body of the paper. Furthermore we assume that

$$\frac{g^{ij}(p)}{|p|^2} \to 0 \, , \text{ as } |p| \to \infty \;\; \forall \, i, j = 1, \dots N \, . \tag{53}$$

We divide the proof in three steps.

Step 1: A is accretive.

By (18),

$$<h|Ah>_{\mathcal{H}} = - <h|Lh>_{\mathcal{H}} + <h|Th>_{\mathcal{H}} = \int g(\partial_p h, \partial_p h) \geq 0 \, .$$

Recall that all integrals are extended over $\mathbb{T}^N \times \mathbb{R}^N$ with measure $dxd\mu$.

Step 2: A is hypoelliptic.

Let $a = \sqrt{g^{-1}}$ (i.e., the positive definite matrix such that $a^2 = g^{-1}$). A straightforward calculation shows that the operator $-A$ can be written in Hörmander's form:

$$-A = \sum_{i=1}^{N} Y_{(i)}^2 + Y_0 \, ,$$

where

$$Y_0 h = (\operatorname{div}_p a) \cdot a \nabla_p h - g(\partial_p E, \partial_p h) - Th \, ,$$
$$Y_{(i)} h = a_i^k \partial_{p^k} h \, .$$

To prove that the operator A is hypoelliptic, we will show that $-A$ satisfies a rank 2 Hormander's condition, namely that the vector fields

$$Y_{(i)} \, , \quad Z_{(i)} := [Y_0, Y_{(i)}]$$

form a basis of $\mathbb{R}^{2N}$. To this purpose we observe that

$$Z_{(i)} = B_i^k \partial_{p^k} + C_i^I \partial_{x^I} = B_i^k P_{(k)} + C_i^I X_{(I)} \, ,$$

where

$$C_i^I = a_i^k \partial_{p^k} v^{(I)}$$

and B is a $p-$dependent $N \times N$ matrix, whose exact form is irrelevant for what follows. Thus the linear transformation $\{X_{(I)}, P_{(i)}\} \to \{Y_{(i)}, Z_{(i)}\}$ is represented by the matrix

$$F = \begin{pmatrix} 0 & a \\ C & B \end{pmatrix} .$$

The determinant of F is given by

$$|\det F| = \det a |\det C| = \det g |\det(\partial_{p^k} v^{(I)})| ,$$

which is positive because $\det g > 0$ and, by Assumption 2, the determinant of the matrix $\partial_{p^k} v^{(I)}$ is non-zero. Thus $\{Y_{(i)}, Z_{(i)}\}$ is a new basis of $\mathbb{R}^{2N}$, concluding the proof.

Step 3: The closure of A is maximally accretive.

By [8, Th. 5.4] (see also [11]), it is enough to prove that the range of $\lambda + A$ is dense in $\mathcal{H}$, for some $\lambda > 0$. We need to show that if $h \in \mathcal{H}$ is such that

$$< h|(\lambda + A) f >_{\mathcal{H}} = 0 , \quad \text{for all } f \in D(A) , \tag{54}$$

then $h = 0$. Note that (54) implies that h is a distributional solution of

$$(\lambda - L - T) h = 0 .$$

Since the operator in the left hand side of the latter equation is hypoelliptic, which is proved as in the previous step, then we may assume that $h \in C^\infty$. Let us begin by proving that the following identity holds:

$$\lambda \int \phi^2 h^2 + \int g(\partial_p(\phi h), \partial_p(\phi h)) = \int h^2 g(\partial_p \phi, \partial_p \phi) - \int h^2 \phi T \phi , \tag{55}$$

for all $\phi \in D(A)$. To prove (55), we use that, by (i)-(ii) of Lemma 1,

$$(\lambda + A)(f_1 f_2) = f_1 (\lambda + A) f_2 + f_2 A f_1 - 2 g(\partial_p f_1, \partial_p f_2) , \quad \text{for all } f_1, f_2 \in C^\infty .$$

Setting $f_1 = \phi$, $f_2 = \phi h$ and multiplying by h the resulting identity we get

$$\phi h (\lambda + A)(\phi h) = h(\lambda + A)(\phi^2 h) - h^2 \phi A \phi + 2 h g(\partial_p \phi, \partial_p(\phi h)) .$$

Integrating and using that $< h|(\lambda + A)(\phi^2 h) >_{\mathcal{H}} = 0$, by (54), we have

$$\int \phi h(\lambda + A)(\phi h) = \int h^2 \phi L \phi - \int h^2 \phi T \phi + 2 \int h g(\partial_p \phi, \partial_p(\phi h)) .$$

Using (18) in the l.h.s. and in the first term in the r.h.s. of the previous identity completes the proof of (55). Now let $k \in \mathbb{N}$ and choose a family of test functions ϕ_k of the form

$$\phi_k(x, p) = \psi(p/k) ,$$

where $\psi \in C_c^\infty$, $0 \le \psi \le 1$, $\psi = 1$ for $p \in B(0, 1/2)$ and $\operatorname{supp} \psi \subset B(0, 1)$. Whence $T\phi_k = 0$. Substituting in (55) we obtain

$$\lambda \int \phi_k^2 h^2 \le \frac{1}{k^2} \int h^2 g^{ij} \partial_{p^i} \psi \partial_{p^j} \psi \, \chi_{|p| \le k} \, .$$

Having assumed (53), we obtain

$$\lambda \int \phi_k^2 h^2 \le \epsilon(k) \, ,$$

where $\epsilon(k) \to 0$ as $k \to \infty$. This finally entails that $h \equiv 0$.

Acknowledgments: This work was completed while the author was a long term participant at the program "Partial Differential Equations in Kinetic Theories" at the Isaac Newton Institute in Cambridge (UK). Support by "Ministerio Ciencia e Innovación", Spain (Project MTM2008-05271) is also acknowledged.

References

[1] A. Arnold, P. Markowich, G. Toscani, A. Unterreiter: On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations. *Comm. Partial Diff. Eqns.* **26** (1&2), 43–100 (2001)

[2] D. Bakry, M. Emery: Hypercontractivité de semi-groupes de diffusion. *C.R. Acad. Sc. Paris.* Série I, Vol. 299 (15), 775–778 (1984)

[3] D. Bakry: *L'hypercontractivité et son utilisation en théorie des semi-groupes.* Lectures Notes in Mathematics 1581, Springer (1994)

[4] J. A. Carrillo, G. Toscani: Exponential convergence toward equilibrium for homogeneous Fokker-Planck type equations. *Math. Meth. Appl. Sci.* **21**, 1269–1286 (1998)

[5] B. Chow, P. Lu, L. Ni: *Hamilton's Ricci flow.* Graduate Studies in Mathematics 77. American Mathematical Society, Providence, RI; Science Press, New York, 2006.

[6] I. Csiszár: Information-type measures of difference of probability distributions. *Stud. Sc. Math. Hung.* **2**, 299–318 (1967)

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

[8] B. Helffer, F. Nier: *Hypoelliptic Estimates and Spectral Theory for the Fokker-Planck Operators and Witten Laplacians.* Lecture Notes in Mathematics 1862, Springer (2005)

[9] F. Hérau, F. Nier: Isotropic hypoellipticity and trend to equilibrium for the Fokker-Planck equation with a high-degree potential. *Arch. Ration. Mech. Anal.* **171**, 151–218 (2004)

[10] L. Hörmander: Pseudodifferential operators and non-elliptic boundary problems. *Ann. of Math.* (2) **83**, 129–209 (1966)

[11] G. Lumer, R. S. Philipps: Dissipative operators in a Banach space. *Pacific J. Math.* **11**, 679–698 (1961)

[12] B. O'Neill: *Semi-Riemannian geometry, with applications to relativity.* Academic Press, Inc. New York (1983)

[13] H. Risken: *The Fokker-Planck equation.* Springer Series in Synergetics Vol. 18, Springer-Verlag, Berlin (1989)

[14] C. Villani: *Hypocoercivity.* Memoirs of the AMS, n. 950, Vol. 202 (2009)