Existence and uniqueness of solutions to Fokker-Planck type equations with irregular coefficients

C. Le Bris[1] & P.-L. Lions[2]

[1]*ENPC, ParisTech, Champs sur Marne, 77455 Marne la Vallée Cedex 2, France*

[2]*Collège de France, 11, place Marcelin Berthelot, 75231 Paris Cedex 05, France, et CEREMADE, Université Paris Dauphine, Place de Lattre de Tassigny, 75775 Paris Cedex 16, France*

CERMICS — ENPC
6 et 8 avenue Blaise Pascal
Cité Descartes - Champs sur Marne
77455 Marne la Vallée Cedex 2

http://cermics.enpc.fr

Existence and uniqueness of solutions to Fokker-Planck type equations with irregular coefficients

C. Le Bris[1] & P.-L. Lions [2]

[1] CERMICS, École Nationale des Ponts et Chaussées,

6 & 8, avenue Blaise Pascal, 77455 Marne-La-Vallée Cedex 2, and

INRIA Rocquencourt, MICMAC project,

Domaine de Voluceau, B.P. 105, 78153 Le Chesnay Cedex, FRANCE

lebris@cermics.enpc.fr

[2] Collège de France, 11, place Marcelin Berthelot,

75231 Paris Cedex 05, FRANCE, and

CEREMADE, Université Paris Dauphine, Place de Lattre de Tassigny,

75775 Paris Cedex 16, FRANCE

lions@ceremade.dauphine.fr

March 26, 2007

Abstract

We study the existence and the uniqueness of the solution to a class of Fokker-Planck type equations with irregular coefficients, more precisely with coefficients in Sobolev spaces $W^{1,p}$. Our arguments are based upon the DiPerna-Lions theory of renormalized solutions to linear transport equations and related equations [6]. The present work extends the results of our previous article [17], where only the simpler case of a Fokker-Planck equation with constant diffusion matrix was addressed. The consequences of the present results on the well-posedness of the associated stochastic differential equations are only outlined here. They will be more thoroughly examined in a forthcoming work [18].

Contents

1 Introduction

We study in this article the existence and the uniqueness of solutions to a large class of Fokker-Planck type equations with irregular coefficients, namely equations of the form

$$\partial_t p - b_i \partial_i p - a_{ij} \partial_{ij}^2 p = 0, \tag{1.1}$$

(or equations of similar forms) with coefficients b_i and a_{ij} that only have Sobolev (typically $W^{1,p}$) regularity. Our work is a follow-up of our previous work [17] where the same equations were considered, in the particular case when a_{ij} is constant. The present work is, like [17], based upon the theory of renormalized solutions and its ingredients, introduced by R. Di Perna and the second author in [6] for linear transport equations (that is, $[a_{ij}] \equiv 0$ in (1.1)). The theory was subsequently extended to other types of equations (such as Boltzmann-type equations) in other works of the same authors.

Because several of our arguments are reminiscent of those of [6] and [17], and because such arguments might require some tedious details, we only focus in the present article on the key issues, and the key manipulations. The reader is spared some unnecessary technicalities. We refer to the previous works for all the detailed arguments. Likewise, for pedagogic purposes, we concentrate on the most illustrative settings. Some specific paragraphs placed after our main arguments and results aim to mention some (among the numerous possible) variants and extensions.

As in [17] (and also in [6]), the well-posedness of the partial differential equation (1.1) has immediate consequences on the well-posedness of the associated differential equations. When $[a_{ij}] \equiv 0$, this differential equation is an ordinary differential equation, see [6] and [17, Section 4]. When $[a_{ij}] \not\equiv 0$, symmetric, nonnegative, the corresponding equations are stochastic differential equations with dispersion matrix σ_{ik} such that $[a_{ij}] = \frac{1}{2}\sigma_{ik}\sigma_{jk}$: see below and also [17, Section 5]. For the sake of brevity, we will postpone our investigation of such questions until a future publication [18]. We concentrate here on all questions related to the partial differential equation *per se*. We however wish to provide the reader of the present article with a complete and self contained view of the relevant issues, and also strongly motivate our choices of mathematical setting. This is why we will mention in Section 3 how the partial differential equations we manipulate here and the stochastic differential equations are intimately connected. We recall some well-known facts and also indicate (which is new) how we intend to proceed and use our results to show the well-posedness of some stochastic differential equations with irregular coefficients.

There is a number of settings where equations of the type (1.1) arise with irregular coefficients, and varying diffusion matrices $[a_{ij}]$ (thus the need to extend our previous results of [17]). Our interest in such theoretical questions stems from the specific setting of complex fluid flows modelling, which we mention in Section 4.

For reasons that will be clear below, and which are precisely related to our motivation regarding the associated stochastic differential equations, we shall mainly concentrate throughout the article on the case when $[a_{ij}]$ is a *symmetric*, *nonnegative*, possibly *degenerate* $N \times N$ matrix, that writes $[a_{ij}] = \dfrac{1}{2}\sigma_{ik}\sigma_{jk}$ for some possibly *non-symmetric* $N \times K$ matrix σ_{ik}. This situation is often the physically relevant one, and is also, as a matter of fact, the most difficult one mathematically. The cases when $[a_{ij}]$ is symmetric definite positive, or symmetric with σ symmetric, are indeed easier. In the former case, the solution is immediately more regular, due to a parabolic regularization effect. In the latter case, some intermediate calculations related to estimates and regularization techniques simplify. Much of the technology developed here is directed towards treating the "difficult" case. See Remark 20 on this matter. On the other hand, the case of a *non-symmetric* matrix $[a_{ij}]$ will not be considered.

Our article is organized as follows. To start with, we emphasize in Section 2 the key mathematical ingredients of our arguments. They come from [6, 17]. We also state there the most illustrative result of the present article. Next, we outline the relation with the differential equations (section 3) and give some motivation (section 4), as announced above. Section 5 is the central part of our work. We show the well-posedness of a particular equation (1.1) that is of divergence form. We first give some formal manipulations that are the bottom line of our arguments, next perform the rigorous proofs. Several variants of our settings and our arguments are possible, and we collect some of them in Section 6. In Section 7, we indicate the necessary modifications to treat other equations of the form (1.1), specifically Fokker-Planck (forward Kolmogorov) equations or backward Kolmogorov equations. Finally, motivated by the comparison with the "classical" setting, we investigate in Section 8 the case of Lipschitz coefficients in (1.1), along with some related issues.

2 Key ingredients and main result

2.1 Transport equations

Our techniques of proof throughout this article are reminiscent of those introduced by R. Di Perna and the second author in [6] (and subsequent articles) in order to establish the existence and uniqueness of solutions to linear transport equations

$$\partial_t f - b_i \partial_i f = 0, \tag{2.1}$$

with Sobolev coefficients. For pedagogic purposes, let us begin by recalling the main ingredients of the approach. The proof falls in two major steps.

The first step consists in establishing formal *a priori* estimates on the tentative solution f. This is performed by multiplying the equation by some function $\beta'(f)$ (where β is some convenient *renormalization* function) and integrating by parts. Formally, this procedure yields

$$\frac{d}{dt}\int \beta(f) + \int (\operatorname{div} \mathbf{b})\beta(f) = 0. \tag{2.2}$$

Therefore, when $\operatorname{div} \mathbf{b}$ is L^∞, we obtain $\int \beta(f)$ bounded for all times if it is bounded at initial time. In particular $\beta(f) = |f|^p$ $(1 \le p < +\infty)$ yields formal L^p bounds on the solution. Besides, the L^∞ bound is obtained by application of the maximum principle. Note that, throughout this section we assume, for simplicity and clarity of exposition, that $\mathbf{b}$ does not depend on time.

Using this first step, existence (which is the "easy" part) is readily proved. The transport coefficient $\mathbf{b}$ is regularized by convolution $\mathbf{b}_\varepsilon = \rho_\varepsilon \star \mathbf{b}$, using some regularizing kernel $\rho_\varepsilon = \varepsilon^{-N} \rho(\varepsilon^{-1}\,\cdot)$, with $\rho \in \mathcal{D}(\mathbb{R}^N)$, $\rho \ge 0$, $\int \rho = 1$. The linear transport equation

$$\partial_t f_\varepsilon - (b_i)_\varepsilon \partial_i f_\varepsilon = 0, \tag{2.3}$$

admits a unique solution f_ε, by standard arguments. The above formal *a priori* estimate (2.2) can be rigorously established on f_ε:

$$\frac{d}{dt} \int \beta(f_\varepsilon) + \int (\operatorname{div} \mathbf{b}) \beta(f_\varepsilon) = 0, \tag{2.4}$$

along with the L^∞ bound. As the equation is linear, passing to the (weak) limit provides a solution to (2.1) in an appropriate functional space, typically $L^1 \cap L^\infty$ when the initial condition $f(t = 0, \cdot)$ lies in that space. For this to hold, we only need $b_i \in L^1_{loc}$ and $\operatorname{div} \mathbf{b} \in L^\infty$. The natural weak formulation of the equation (2.1) also readily follows from the above argument.

The second major step is a regularization procedure. It is based upon the so-called *commutation* lemma, which basically claims that

$$[\rho_\varepsilon, \mathbf{b} \cdot \nabla](f) := \rho_\varepsilon \star (\mathbf{b} \cdot \nabla f) - \mathbf{b} \cdot \nabla(\rho_\varepsilon \star f) \overset{\varepsilon \to 0}{\longrightarrow} 0, \tag{2.5}$$

in L^1_{loc} when $\mathbf{b} \in W^{1,1}_{loc}$ and $f \in L^\infty$ (for instance). We refer to Lemma 1 below for a precise statement. The need for some Sobolev regularity on $\mathbf{b}$ may be formally understood in the following manner: the above commutator basically involves a quantity of the form

$$\int f(y)\, (\mathbf{b}(y) - \mathbf{b}(x)) \cdot \nabla \rho_\varepsilon(x - y)\, dy \tag{2.6}$$

where ρ_ε converges in distribution to the Dirac mass, thus the need for evaluating $\mathbf{b}(y) - \mathbf{b}(x)$ in terms of $y - x$ for $y - x$ small. Uniqueness readily follows from this commutation lemma by convolution: considering $f = g - h$ the difference of two solutions to (2.1), one convoluates the transport equation (2.1) with ρ_ε, next obtain the same equation set on f_ε up to an error term:

$$\partial_t f_\varepsilon - b_i \partial_i f_\varepsilon = [\rho_\varepsilon, \mathbf{b} \cdot \nabla](f), \tag{2.7}$$

We multiply both sides by $\beta'(f_\varepsilon)$ and integrate over the whole space. Letting next ε go to zero, (2.2) is obtained, which immediately gives $f = 0$ (because it is zero at initial time). Thus uniqueness holds. In the above argument, some cut-off functions are needed, for the integrals over the whole space to be conveniently treated. Boundary terms are then taken care of using appropriate growth conditions at infinity for $\mathbf{b}$. We however omit such technicalities in the present oversimplified outline.

A summary of the above argument could be as follows. The main two ingredients are

- an *a priori* estimate,
- a regularization procedure.

For these two steps to be possible,

- some global bounds on $\operatorname{div}\mathbf{b}$ (typically $\operatorname{div}\mathbf{b} \in L^\infty$) are necessary for a priori estimates to hold,
- some local, typically $W^{1,p}$, regularity on $\mathbf{b}$ is needed for regularization,
- in addition to this, some technical assumptions are needed, regarding growth at infinity on $\mathbf{b}$ (typically $\dfrac{\mathbf{b}}{1+|x|} \in L^1 + L^\infty$) for integration by parts over the whole space to be rigorously performed.

The intuitive belief follows that, whenever the above two essential steps (*a priori* estimate and regularization) may be carried out, existence and uniqueness hold. Several results in the vein of [6], including [17] and the present work, show that this heuristic belief indeed holds true.

2.2 Fokker-Planck equations

In [17], we have studied the parabolic equation

$$\partial_t f - b_i \partial_i f - \frac{1}{2}\Delta f = 0. \tag{2.8}$$

Following the above two steps, it is immediately seen that the presence of the regularizing second order operator $-\Delta$ yields a better regularity on the solution f than in the pure transport case. Indeed, the formal *a priori* estimate obtained from $\beta(f) = \frac{f^2}{2}$ reads

$$\frac{d}{dt}\int \frac{f^2}{2} + \int (\operatorname{div}\mathbf{b})\frac{f^2}{2} + \int |\nabla f|^2 = 0. \tag{2.9}$$

Some $L^2([0,T], H^1)$ bound on f is deduced. This H^1 regularity can be in turn used in the regularization step. The term (2.6) indeed writes (up to a term in $\operatorname{div}\mathbf{b}$)

$$\int (\mathbf{b}(y) - \mathbf{b}(x)) \cdot \nabla f(y)\, \rho_\varepsilon(x-y)\, dy. \tag{2.10}$$

We see that a L^2 regularity on $\mathbf{b}$ is now sufficient to proceed, without the need for a $W^{1,p}$ regularity. The details are worked out in [17, Section 5], where we proved that existence and uniqueness of the solution to the equation (2.8) holds in $L^\infty([0,T], L^1 \cap L^\infty) \cap L^2([0,T], H^1)$ when $\mathbf{b} = \mathbf{b}^1 + \mathbf{b}^2$, $\mathbf{b}^1 \in W^{1,1}_{loc}$, $\mathbf{b}^2 \in L^2_{loc}$, $\operatorname{div} \mathbf{b} \in L^\infty$ and $\dfrac{\mathbf{b}}{1+|x|} \in L^1 + L^\infty$.

Remark 1 *Some variants and extensions are also indicated in the same work. Alternately, an extension of the standard equation (2.1) in another direction is also examined in [17]. It is the specific case of the transport equation*

$$\partial_t f(t,x_1,x_2) - \mathbf{b}^1(x_1) \cdot \nabla_{x_1} f(t,x_1,x_2) - \mathbf{b}^2(x_1,x_2) \cdot \nabla_{x_2} f(t,x_1,x_2) = 0, \quad (2.11)$$

with $\mathbf{b}^1 \in W^{1,1}_{x_1}$ *and* $\mathbf{b}^2 \in L^1_{x_1}(W^{1,1}_{x_2})$. *In other words we only have* partial $W^{1,1}$ *regularity of* $\mathbf{b}^2$ *in the* x_1 *variable. Some consequences of this result on the theory of generalized flows for ordinary differential equations are presented in [17, Section 4]. In particular, the relevance for a specific physical problem is mentionned there. We will return to this in Section 4.*

The purpose of the present work is to more thorougly investigate the application of the above methodology to parabolic type equations. The immediate generalization of (2.8) is

$$\partial_t f - b_i \partial_i f - \frac{1}{2} \sigma_{ik} \sigma_{jk} \partial^2_{ij} f = 0 \quad (2.12)$$

corresponding to a varying dispersion matrix σ, while (2.8) corresponds to a constant dispersion matrix σ. More generally, the case of several variants of (1.1) will be examined.

On the basis of the above outline, it is immediate to realize that H^1 regularity is expected on the solution f as soon as the second order differential operator $-\dfrac{1}{2}\sigma_{ik}\sigma_{jk}\partial^2_{ij}$ is positive definite. This will be the purpose of Section 6.3. Even if the operator is only nonnegative and possibly degenerate, we will see that we can still proceed, and find some relatively general setting for (2.12) to be well-posed.

The following result exemplifies many results in this direction. It is a particular case of the general Proposition 5, proved in Section 7.

From Proposition 5 [Section 7 of the present work]

Set $b^\sigma_i = b_i - \frac{1}{2}\partial_j\left(\sigma_{ik}\sigma_{jk}\right)$ *for all* $1 \le i \le N$. *Assume* $\mathbf{b}^\sigma$ *and* σ *are time-independent and satisfy:*

$$\mathbf{b}^\sigma \in \left(W^{1,1}_{loc}(\mathbb{R}^N)\right)^N, \quad \operatorname{div} \mathbf{b}^\sigma \in L^\infty(\mathbb{R}^N), \quad \frac{\mathbf{b}^\sigma}{1+|x|} \in \left(L^1 + L^\infty(\mathbb{R}^N)\right)^N$$

$$\sigma \in \left(W^{1,2}_{loc}(\mathbb{R}^N)\right)^{N\times K}, \quad \frac{\sigma}{1+|x|} \in \left(L^2 + L^\infty(\mathbb{R}^N)\right)^{N\times K}.$$

Then, for each initial condition in $L^2 \cap L^\infty$, *equation (2.12), that is*

$$\partial_t f - b_i \partial_i f - \frac{1}{2} \sigma_{ik} \sigma_{jk} \partial^2_{ij} f = 0$$

has a unique solution in the space

$$\left\{ f \in L^\infty([0,T], L^2 \cap L^\infty), \quad \sigma^t \nabla f \in L^2([0,T], L^2) \right\}.$$

Several variants and extensions of the above result will also be considered throughout the present article.

The consequence of Proposition 5 on the well-posedness of some stochastic differential equations with irregular coefficients will hardly be approached here. As announced above, it is the specific purpose of the companion article [18]. Anticipating on [18], we however would like to briefly comment on such issues, for they motivate the present work and have important connections with the specific settings we choose for developping our arguments. This is the purpose of the next section.

To close this section (and slightly anticipating on the contents of the next one), let us mention some, important, related works. The result by Di Perna and the second author on existence and uniqueness for transport equations for Sobolev regular vector fields has been generalized in a series of works by L. Ambrosio and collaborators to BV vector fields (see [1, 2, 3] for various aspects). In the same vein, the result mentionned above on transport equations of the particular form (2.11) has also been generalized to BV vector fields by N. Lerner [16]. As regards parabolic type equations like (2.12), a recent work[1] by A. Figalli [10] shows existence and uniqueness for two main settings: a) uniformly definite positive, lipschitz regular in time, matrices $\sigma\sigma^T$ and L^∞ vector fields, or b) space-independent matrices $\sigma\sigma^T$ and BV vector fields. Other assumptions (like growth at infinity and control of the divergence) are basically similar to those above. Interestingly, in all these works, when coming to the well-posedness of the associated (ordinary or stochastic) differential equations, the viewpoint adopted is somewhat different from that of the works [6, 17]: the connection is performed through the introduction of an adequate measure on the space of paths (see the works mentioned above for more details). For obvious reasons, this viewpoint, alternate from ours, seems to be particularly convenient for the adaptation to the stochastic setting and the notion of martingale solutions, as exemplified in [10]. We shall return to this in [18].

On the other hand, the adaptation of the specific results contained in the present work to the case of BV vector fields should be an interesting, amenable, extension.

3 Relation with differential equations

3.1 Ordinary differential equations

It is well known that, in the case of regular coefficients $\mathbf{b}$ (which we again for simplicity consider time-independent), solving the linear transport equation

[1]The authors are grateful to A. Figalli for communicating a preprint version of [10].

(2.1)

$$\partial_t f - b_i \partial_i f = 0$$

is closely related to solving for all initial conditions x the ordinary differential equation:

$$\begin{cases} \dot{X} = \mathbf{b}\,(X), \\ X(t=0) = x. \end{cases} \tag{3.1}$$

More precisely, the connection between the partial differential equation (2.1) and the differential equation (3.1) is ensured by the method of characteristics, also called method of lines. When $X(t, x)$ denotes the solution to (3.1), then the solution to (2.1) starting from the initial condition f_0 at time $t = 0$ reads

$$f(t, x) = f_0(X(t, x)). \tag{3.2}$$

Conversely, successively solving the transport equation (2.1) with as initial condition all coordinates fields provides the solution X to (3.1).

This correspondence for regular fields $\mathbf{b}$ was extended by R. Di Perna and the second author in [6] in order to define a notion of generalized flow of solutions for ordinary differential equations with Sobolev coefficients. To cut a long story short, it was proved in [6] that for $\mathbf{b} \in W^{1,1}$, $\operatorname{div}\mathbf{b} \in L^\infty$, $\dfrac{\mathbf{b}}{1+|x|} \in \left(L^1 + L^\infty\right)^N$, one may define a solution flow for (3.1, precisely because, under the same assumptions, one may solve (2.1). Furthermore, it was proved in [19] that for all L^1 fields $\mathbf{b}$ (satisfying the other two condition above), the existence and uniqueness of the generalized flow is equivalent to the existence and uniqueness of the solution to the transport equation. Any additional property that ensures one of the two facts then implies the other fact.

Remark 2 *Let us mention that some specific remarks and extensions on the above questions will be the subject of [13].*

3.2 Stochastic differential equations

A similar line of thought may be followed for stochastic differential equations. Our intent is only to provide the reader not familiar to the field with a rapid introduction of the key connections between second order partial differential equations and stochastic differential equations . We therefore omit here all technicalities related to the rigorous setting of stochastic differential equations, and proceed somewhat formally. We refer to the excellent textbooks [14, 15, 23, 24, 25] for all the very important mathematical details of all the aspects adressed here. Of course, the present section is no more than a brief, convenient, surrogate to such treatises.

3.2.1 Regular coefficients

It is well known that there are two notions of solutions to the stochastic differential equation

$$d\mathbf{X}_t = \mathbf{b}\,(\mathbf{X}_t)\,dt + \sigma(\mathbf{X}_t)\,d\mathbf{W}_t. \tag{3.3}$$

The solution is *strong* if it exists for a given probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, a given brownian motion $\mathbf{W}_t$, and a given initial condition $\mathbf{X}_0$. It is said *weak* if $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, $\mathbf{W}_t$, and the law of $\mathbf{X}_0$ are part of the solution.

For strong solutions, the right notion of uniqueness is *pathwise uniqueness* (that states the uniqueness of $\mathbf{X}_t$ for $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, $\mathbf{W}_t$, $\mathbf{X}_0$ given). For weak solutions, even if pathwise uniqueness may be defined in an analogous manner, a more appropriate notion of uniqueness is *uniqueness-in-law*, stating that any two weak solutions sharing the same initial distribution at initial time have the same law for all times.

Additionnaly, under appropriate regularity conditions on the coefficients, the existence of a weak solution to the equation is equivalent to the existence of a solution to the *Martingale problem* of Stroock and Varadhan. The uniqueness in law of the weak solution is also equivalent to the unique solvability of the Martingale problem for any fixed initial distribution.

To illustrate all the above, let us recall the following two prototypical results of existence and uniqueness, for regular coefficients $\mathbf{b}$ and σ.

Theorem 1 from [15, p 289] *Suppose that* $\mathbf{b}$ *and* σ *satisfy the growth condition*

$$\frac{\mathbf{b}}{1+|x|} \in L^\infty(\mathbb{R}^N), \tag{3.4}$$

and

$$\frac{\sigma(x)}{1+|x|} \in \left(L^\infty(\mathbb{R}^N)\right)^{N\times K}. \tag{3.5}$$

Suppose they satisfy the global Lipschitz condition

$$\left\|\mathbf{b}\,(x) - \mathbf{b}\,(y)\right\|_{\mathbb{R}^N} + \left\|\sigma(x) - \sigma(y)\right\|_{\mathbb{R}^{N\times K}} \le K \left\|x - y\right\|_{\mathbb{R}^N} \tag{3.6}$$

Then there exists a unique (strong) solution to the stochastic differential equation (5.1).

Theorem 2 [25, p170] *Suppose* $\mathbf{b}$ *is measurable and satisfies the growth condition*

$$\frac{\mathbf{b}}{1+|x|} \in L^\infty(\mathbb{R}^N). \tag{3.7}$$

Suppose in addition that $\sigma(x)\sigma^t(x)$ *is continuous, positive definite for all* $x \in \mathbb{R}^N$*, and satisfies the growth condition*

$$\frac{\sigma(x)\sigma^t(x)}{1+|x|^2} \in \left(L^\infty(\mathbb{R}^N)\right)^{N\times N}. \tag{3.8}$$

Then there exists a unique solution to the martingale problem.

The relation to partial differential equations of the form (1.1) may now be outlined. Two basic facts need to be recalled. They will be useful to understand our motivation in the sequel.

When $\mathbf{b}$ and σ are lipschitz (more precisely satisfy the assumptions of Theorem 1 above), a strong solution to (3.3) exists, pathwise unique. Besides, a solution to

$$\partial_t f - b_i \partial_i f - \frac{1}{2}\sigma_{ik}\sigma_{jk}\partial^2_{ij} f = 0$$

uniquely exists for all continuous initial condition f_0. The notion of solution is that of viscosity solutions, and the solution is continuous for all times. The Feynmann-Kac representation formula (essentially a generalization to (3.2))

$$f(t,x) = \mathbb{E}\left(f_0(X_t(x))\right) \tag{3.9}$$

gives the connection between the two viewpoints. On the other hand, weak solutions may also be understood in terms of a partial differential equation. The most illustrative connection is that a weak solution is unique in law as soon as one may uniquely solve the Fokker-Planck equation

$$\partial_t p - \operatorname{div}(p\mathbf{b}) - \frac{1}{2}\partial^2_{ij}\left(\sigma_{ik}\sigma_{jk}p\right) = 0. \tag{3.10}$$

In addition, the existence of a solution to

$$\partial_t f - b_i \partial_i f - \frac{1}{2}\sigma_{ik}\sigma_{jk}\partial^2_{ij} f = 0$$

in an appropriate functional space implies the uniqueness of the solution of the Martingale problem for each initial condition.

3.2.2 Irregular coefficients

In [17], we extended the connection between uniqueness in law for (3.3) and uniqueness for (3.10) to the case of a *constant* dispersion matrix σ and a Sobolev drift vector $\mathbf{b}$. More precisely, we proved that for $\sigma \equiv \mathbb{I}\mathrm{d}$ and, again, $\mathbf{b} \in \left(W^{1,1}_{loc}\right)^N$, $\operatorname{div}\mathbf{b} \in L^\infty$, $\dfrac{\mathbf{b}}{1+|x|} \in \left(L^1 + L^\infty\right)^N$,

- we can uniquely solve the Fokker-Planck equation (3.10),
- we may define a generalized flow of solutions to (3.3), strong in the probability sense (which amounts to solving the SDE for almost all initial conditions),
- these strong solution flows all share the same law, which is the unique solution to (3.10).

In addition, we proved that we may also uniquely solve the Fokker-Planck equation (3.10) for $\mathbf{b} \in L^2 + W^{1,1}$ (along with the other two conditions). We however pointed out in [17, p128] that for such a drift vector $\mathbf{b}$ the notion of generalized solutions to (3.3) was to be extended before we are in position to claim that, in that case, we have a unique-in-law solution to (3.3), weak in the sense of probability and generalized in the sense of the Di Perna Lions theory. The reason for this (temporary) obstruction is related to our very method of proof in [17]. We circumvent the difficulty of directly defining a generalized flow for stochastic differential equations specifically using that, when $\sigma \equiv \mathbb{I}\mathrm{d}$, the stochastic differential equation can be recasted as an ordinary differential equation parameterized by the randomness (see the discussion below). Correspondingly, the generalized flow for the stochastic differential equation is obtained from that of the ordinary differential equation. Now the latter only exists when $\mathbf{b} \in W^{1,1}$. On the other hand, when $\mathbf{b} \in L^2$, only a direct strategy, specifically using the regularizing nature of the stochastic term, may allow to define the flow. This goes beyond the simple case $\sigma \equiv \mathbb{I}\mathrm{d}$ and is not completed in [17].

To state it otherwise, the following questions remained unsolved in [17]:

- in the case of constant dispersion matrix σ,
  - the pathwise uniqueness of (generalized) strong solutions to (3.3) when $\mathbf{b}$ has Sobolev regularity,
  - the existence and uniqueness-in-law of (generalized) weak solutions to (3.3) when $\mathbf{b}$ is, say, L^2,
- in the case of varying dispersion matrix σ: all questions related to existence and uniqueness.

The purpose of the work [18] will be to address such issues using the results of the present article on Fokker Planck type equations.

For questions related to pathwise uniqueness, obviously a more demanding issue than uniqueness-in-law, we will make use of the Fokker-Planck equation in a space of doubled dimension. This important tool is briefly reviewed in the following section.

For questions related to a varying dispersion matrix σ, we will have to completely revisit the approach adopted in [17]. It is intuitively easy to see why the case of a varying dispersion matrix σ is significantly more difficult that the case when it is constant, as regards uniqueness issues. The following formal arguments hopefully illustrate this. As is well known, proving uniqueness is basically substracting one equation to the other. In the case when σ is constant $d\mathbf{X}_t = \mathbf{b}(\mathbf{X}_t)\,dt + \sigma\,d\mathbf{W}_t$ and $d\mathbf{Y}_t = \mathbf{b}(\mathbf{Y}_t)\,dt + \sigma\,d\mathbf{W}_t$ yield $d(\mathbf{X}_t - \mathbf{Y}_t) = (\mathbf{b}(\mathbf{X}_t) - \mathbf{b}(\mathbf{Y}_t))\,dt$. The brownian motion cancels out and we are left with an *ordinary* differential equation set on stochastic processes. In other terms, proving uniqueness is studying the linear tangent operator, which is a simple object when σ is constant. The same argument, again put differently, consists in writing $d(\mathbf{X}_t - \sigma\,\mathbf{W}_t) = \mathbf{b}\left((\mathbf{X}_t - \sigma\,\mathbf{W}_t) + \sigma\,\mathbf{W}_t\right)\,dt$ instead of $d\mathbf{X}_t = \mathbf{b}(\mathbf{X}_t)\,dt + \sigma\,d\mathbf{W}_t$. We explicitly use this in [17] to define the generalized flows of stochastic differential equations with constant σ from the existing generalized flows of ordinary differential equations. All previous arguments unfortunately collapse when σ varies. The previous computation gives

$d(\mathbf{X}_t - \mathbf{Y}_t) = \mathbf{b}\,(\mathbf{X}_t) - \mathbf{b}\,(\mathbf{Y}_t))\,dt + (\sigma(\mathbf{X}_t) - \sigma(\mathbf{Y}_t))\,d\mathbf{W}_t$. Likewise, the connection between the stochastic differential equation and the ordinary differential equation is unclear. A last facet of the above striking difference that we would like to emphasize is the following. Generalized flows correspond to solutions defined for almost all initial conditions. Proving their existence amounts to ensuring that, as time advances, the set of singular initial conditions remains of Lebesgue measure zero. When σ is constant, the brownian motion does not significantly perturb the evolution of such singularities. It simply "shifts" trajectories. On the other hand, a dispersion matrix $\sigma(\mathbf{X}_t)$ strongly modifies trajectories, and this modification has to be understood and controlled.

3.3 A tool for pathwise uniqueness

One important ingredient of our forthcoming work [18] is the consideration of a Fokker-Planck type equation in a space of doubled dimension. This will allow us to prove pathwise uniqueness. Let us briefly mention here our strategy.

It is well known, and has been recalled above, that the unique solvability of the Fokker-Planck equation implies the uniqueness-in-law of the solution of the stochastic differential equation. This is standard for regular coefficients. The case of less regular drift vectors (with $\sigma \equiv \mathrm{Id}$ though) has also been mentioned above.

We now indicate a possible way for establishing pathwise uniqueness, which allows for generalizations to less regular coefficients. In order to prove pathwise uniqueness for (3.3)

$$d\mathbf{X}_t = \mathbf{b}\,(\mathbf{X}_t)\,dt + \sigma(\mathbf{X}_t)\,d\mathbf{W}_t,$$

it is possible to consider the following Fokker-Planck equation

$$\begin{aligned}&\frac{\partial p(t,x,y)}{\partial t} + \mathrm{div}_{\,x}(p(t,x,y)\mathbf{b}\,(x)) + \mathrm{div}_{\,y}(p(t,x,y)\mathbf{b}\,(y))\\ &-\frac{1}{2}\frac{\partial^2}{\partial x_i \partial x_j}(\sigma_{ik}(x)\sigma_{jk}(x)\,p) - \frac{\partial^2}{\partial x_i \partial y_j}(\sigma_{ik}(x)\sigma_{jk}(y)\,p)\\ &-\frac{1}{2}\frac{\partial^2}{\partial y_i \partial y_j}(\sigma_{ik}(y)\sigma_{jk}(y)\,p) = 0,\end{aligned} \tag{3.11}$$

with solution $p = p(t,x,y)$. This equation is of course the Fokker Planck equation associated with the stochastic differential equation

$$d\begin{pmatrix}\mathbf{X}_t\\ \mathbf{Y}_t\end{pmatrix} = \mathbf{B}(\mathbf{X}_t,\mathbf{Y}_t) + \mathbf{\Sigma}(\mathbf{X}_t,\mathbf{Y}_t)d\begin{pmatrix}\mathbf{W}_t\\ \mathbf{W}_t\end{pmatrix} \tag{3.12}$$

set in $\mathbb{R}^{2N}$ while (3.3) is set in $\mathbb{R}^N$, with coefficients

$$\mathbf{B}(x,y) = \begin{pmatrix}\mathbf{b}\,(x)\\ \mathbf{b}\,(y)\end{pmatrix} \qquad \text{and} \quad \mathbf{\Sigma}(x,y) = \begin{pmatrix}\sigma(x) & 0\\ \sigma(y) & 0\end{pmatrix} \tag{3.13}$$

Likewise, (3.11) can be written in the more compact form

$$\frac{\partial p}{\partial t} + \mathrm{div}_{\,x,y}(p\mathbf{B}(x,y)) - \frac{1}{2}D^2(\mathbf{\Sigma}(x,y)\mathbf{\Sigma}^t(x,y)p) = 0 \tag{3.14}$$

with $\mathbf{B}(x,y)$ as in (3.13) and

$$\mathbf{\Sigma}\mathbf{\Sigma}^t = \begin{pmatrix} \sigma(x)\sigma^t(x) & \sigma(x)\sigma^t(y) \\ \sigma(y)\sigma^t(x) & \sigma(y)\sigma^t(y) \end{pmatrix} \tag{3.15}$$

As it stands, the double Fokker-Planck equation (3.11) appears of course as a particular case of a standard Fokker-Planck equation.

Note, which will be important for the sequel, that its dispersion matrix $\mathbf{\Sigma}$ given by (3.13) is by construction non-symmetric and that $\mathbf{\Sigma}\mathbf{\Sigma}^t$ in (3.15) is, by construction again, degenerate. This is the main *mathematical* reason why we mainly concentrate here on symmetric possibly degenerate diffusion matrices coming from non-symmetric matrices σ (only considering full rank dispersion matrices or symmetric σ for the sake of comparison). Our arguments and results will then readily apply to the double Fokker-Planck equation, and thus open the way to pathwise uniqueness results. In fact, there is another reason for mainly considering degenerate diffusion matrices, related to the specific physical context that originally motivates this work. We want our setting to cover the classical (that is, non doubled!) Fokker-Planck equation for some flows of complex fluids where, indeed, the diffusion is degenerate. This will be seen in Section 4.

Arguments based on such types of Fokker-Planck equation in a space of doubled dimension are well known (see [25, p 198]). They have already been used in [12] and [4], for purposes similar to those of the present work.

We are now in position to briefly sketch why uniqueness of the solution to (3.14) may imply pathwise uniqueness for the original stochastic differential equation (3.3). We simply remark that for $\mathbf{X}_t(x)$ and $\mathbf{Y}_t(y)$ two solutions to (3.3) (respectively starting from x and y), the joint law of $(\mathbf{X}_t(x), \mathbf{Y}_t(y))$ solves (3.14). So does the joint law of $(\mathbf{X}_t(x), \mathbf{X}_t(y))$. By uniqueness, these two laws are therefore equal to one another. Formally taking the limit $y \longrightarrow x$, we obtain $\mathbf{X}_t(x) = \mathbf{Y}_t(x)$ and thus pathwise uniqueness.

Alternately, one may use the adjoint viewpoint. Instead of the Fokker-Planck equation (3.11), one may consider its adjoint form, namely the backward Kolmogorov equation:

$$\begin{aligned} &\frac{\partial f}{\partial t} - b_i(x)\frac{\partial f}{\partial x_i} - b_i(y)\frac{\partial f}{\partial y_i} \\ &\quad -\frac{1}{2}\sigma_{ik}(x)\sigma_{jk}(x)\frac{\partial^2 f}{\partial x_i \partial x_j} - \sigma_{ik}(x)\sigma_{jk}(y)\frac{\partial^2 f}{\partial x_i \partial y_j} \\ &\quad -\frac{1}{2}\sigma_{ik}(y)\sigma_{jk}(y)\frac{\partial^2 f}{\partial y_i \partial y_j} = 0 \end{aligned} \tag{3.16}$$

with solution $f = f(t,x,y)$. Owing to the Feynmann-Kac formula (that is the stochastic version of (3.2)), the function $f(t,x,y) = \mathbb{E}(|\mathbf{X}_t(x) - \mathbf{Y}_t(y)|^2)$ solves equation (3.16) with initial condition $\psi_0(x, y) = |x-y|^2$, when $\mathbf{X}_t(x)$ and $\mathbf{Y}_t(y)$ are again two solutions to (3.3), starting from x and y respectively. A simple computation then shows that, for some suffificently large constant C (depending on the Lipschitz continuity of $\mathbf{b}$ and σ, a property which implies the uniqueness of the solution to (3.16)), the function $\overline{\psi}(t, x, y) = e^{Ct}|x-y|^2$ is a supersolution to (3.16). It follows that $\mathbb{E}(|\mathbf{X}_t(x)-\mathbf{Y}_t(y)|^2)) \leq e^{Ct}|x-y|^2$ for all times. Taking $x = y$ yields pathwise uniqueness.

Both arguments (on (3.14) and on (3.16)) of course hold for *regular* (say Lipschitz) coefficients $\mathbf{b}$ and σ. The purpose of our work [18] will be to try and extend the arguments so as to treat "general" coefficients $\mathbf{b}$ and σ. Depending on our assumptions on $\mathbf{b}$ and σ, existence and uniqueness of generalized stochastic flows for stochastic differential equations (3.3) will be studied, either in law (arguing on (3.14)), or pathwise (arguing on (3.10)).

4 Motivation: Modelling of polymeric fluids

As mentioned in the introduction, there are many situations when stochastic differential equations (and correspondingly Fokker-Planck type equations) arise with irregular drift vectors $\mathbf{b}$ and dispersion matrix σ. Fluid mechanics is a field where this is often the case. The specific application that has originally motivated our work is the multiscale modelling of complex fluid flows.

Our first contribution [17] was already motivated by the so-called micro-macro modelling of infinitely dilute solutions of *flexible* polymers. In such fluids, the macroscopic equations of conservation of mass and momentum are coupled with a kinetic mesoscopic description of the evolution of the microstructures. The macroscopic velocity field, which has no particular reason to be regular (and which, in any event, is not regular at the discretization level because it is typically a finite element approximation) enters as a parameter the stochastic differential equation or the Fokker-Planck equation describing the evolution of the microstructure. In the case of *flexible* polymers, a typical equation arising in this context reads

$$d\mathbf{R}_t + u \cdot \nabla_x \mathbf{R}_t = (\nabla u \mathbf{R}_t - \nabla V(\mathbf{R}_t))\, dt + d\mathbf{W}_t. \tag{4.1}$$

In (4.1), u denotes the macroscopic velocity of the fluid, x is the macroscopic space variable, $\mathbf{R}_t$ is a vector-valued random variable that is implicitly indexed by the macroscopic variable x and describes the statistical state of the microstructure at macropoint x. It is typically the end-to-end vector of the polymeric chain. The term $u \cdot \nabla_x \mathbf{R}_t$ models the transport of the microstructures by the fluid, the term $\nabla u \mathbf{R}_t$ models their elongation, the potential V is a potential of entropic nature, and $d\mathbf{W}_t$ is a brownian motion that models the permanent collisions of the chain with the solvent molecules. More details on this model and on those we will briefly introduce below may be found in the excellent (mathematically oriented) textbook [21] and the recent monography [22].

In [17], we explained how a rigorous meaning can be given to (4.1) even though the velocity field u is not regular. The bottom line of our argument is to simultaneously integrate (4.1) along the characteristic lines, in order to eliminate the transport term $u \cdot \nabla_x \mathbf{R}_t$ (we thus adopt a Lagrangian viewpoint), and also handle the right-hand side of (4.1) as a stochastic differential equation with "irregular" coefficients. In fact, we consider the system formed by (4.1) and

$$\left\{ \begin{array}{lcl} d\,X(t) & = & u(X(t))\,dt, \\ d\mathbf{R}_t(X(t)) & = & (\nabla u \mathbf{R}_t(X(t)) - \nabla V(\mathbf{R}_t(X(t))))\;dt + d\mathbf{W}_t \end{array} \right. \tag{4.2}$$

as a mixed stochastic/deterministic equation that we treat as a whole. The strategy for giving a meaning to this system is then a combination of the existence and uniquess result for a transport equation of the "mixed" form (2.11) and the existence and uniqueness result for Fokker-Planck equations with constant diffusion matrix (2.8). The relevant partial differential equation, written with obvious notation, is

$$\partial_t\, f(t,x,r) - \mathbf{b}^{\,x}(t,x)\cdot\nabla_x f(t,x,r) - \mathbf{b}^{\,r}(t,x,r)\cdot\nabla_r f(t,x,r) - \frac{1}{2}\Delta_r f(t,x,r) = 0.$$

As the strategy above does not pose any specific problem once the key ingredients are present, we have not presented any detailed argument in [17].

Unfortunately the above result does not cover all physically relevant cases. There exists *e.g.* a category of polymeric fluids that, as opposed to flexible polymers, are more appropriately modelled by *rigid rods.* For such polymers, (4.1) is to be replaced (see [21, p 249]) by

$$\begin{aligned} d\mathbf{X}_t \;\; = \;\; & \left(\left(\mathrm{I\!Id} - \frac{\mathbf{X}_t \otimes \mathbf{X}_t}{\|\mathbf{X}_t\|^2}\right) \mathbf{A}(t,\mathbf{X}_t) \; - \frac{N-1}{2}\, B^2 \frac{\mathbf{X}_t}{\|\mathbf{X}_t\|^2} \right) dt \\ & + B \left(\mathrm{I\!Id} - \frac{\mathbf{X}_t \otimes \mathbf{X}_t}{\|\mathbf{X}_t\|^2}\right) d\mathbf{W}_t. \end{aligned} \tag{4.3}$$

For simplicity and in order to concentrate on the crucial difference with (4.1) we have omitted the transport term in the left-hand side. In (4.3), $\mathbf{X}_t$ denotes the rigid-rod vector, $\mathrm{I\!Id}$ denotes the Identity matrix of size N, B is a fixed parameter, $d\mathbf{W}_t$ is a N dimensional Brownian motion, and $\mathbf{A}$ is the vector of size N defined by

$$\mathbf{A}(t,\mathbf{X}_t) = \nabla u(t,y)\,\mathbf{X}_t - \nabla V(\mathbf{X}_t), \tag{4.4}$$

As in (4.1), V is the potential of the entropic force the polymer is subjected to. It is immediate to understand that the projection operator $\mathrm{I\!Id} - \dfrac{\mathbf{X}_t \otimes \mathbf{X}_t}{\|\mathbf{X}_t\|^2}$ aims to preserve the length of $\mathbf{X}_t$ in time (thus the "rigidity" of the rod).

The major difference with (4.1) is that the dispersion matrix σ of (4.3) is no longer a constant. The existence of a varying dispersion matrix $\sigma = \sigma(\mathbf{X}_t)$ motivates the whole present study. In addition, for all $\zeta \in \mathbb{R}^N$,

$$\frac{1}{B^2}\,(\sigma\sigma^t\zeta,\zeta) = \frac{\|\mathbf{x}\|^2\|\zeta\|^2 - (\mathbf{x},\zeta)^2}{\|\mathbf{x}\|^2},$$

where $(\cdot,\cdot)$ denotes the euclidean scalar product on $\mathbb{R}^N$ and $\|\cdot\|$ the associated norm. It follows that $\sigma\sigma^t$ cannot be positive definite, whatever $\mathbf{x} \in \mathbb{R}^N$. This is a second motivation for focusing on possibly degenerate matrices, in addition to the point made above on the structure of the doubled Fokker-Planck equation.

Of course, the above dispersion matrix $\sigma(\mathbf{X}_t)$ is to be considered for normalized $\mathbf{X}_t$ (or $\mathbf{X}_t$ close to be of unit norm) since equation (4.3) propagates forward in time the normalization. In the vicinity of such normalized $\mathbf{X}_t$, the dispersion matrix is a regular function of $\mathbf{X}_t$. The techniques of the present article, although needed for the irregular drift in (4.3), are thus not *specifically* required for the dispersion. This is however an interesting case to consider, perhaps thinking to other contexts.

To conclude with these modelling issues, let us mention that the full generality of physically relevant cases is not covered yet even when considering varying dispersion matrices $\sigma = \sigma(\mathbf{X}_t)$. Another case of practical interest, in the spirit of the modelling seen above, is the modelling of *liquid crystal* polymers (see [21, p 253]). Basically, the stochastic differential equation involved in such a modelling reads:

$$\begin{aligned} d\mathbf{X}_t = \Bigg[\Big(\mathbb{I}\mathrm{d} - \frac{\mathbf{X}_t \otimes \mathbf{X}_t}{\|\mathbf{X}_t\|^2}\Big) \Big(\nabla u(t,y)\, \mathbf{X}_t - \frac{1}{2} B^2(\mathbf{X}_t) \nabla V_e(\mathbf{X}_t) \\ + \frac{1}{2}\nabla(B^2)(\mathbf{X}_t)\Big) - B^2(\mathbf{X}_t)\, \frac{\mathbf{X}_t}{\|\mathbf{X}_t\|^2} \Bigg] dt \qquad (4.5) \\ + B(\mathbf{X}_t)\Big(\mathbb{I}\mathrm{d} - \frac{\mathbf{X}_t \otimes \mathbf{X}_t}{\|\mathbf{X}_t\|^2}\Big)\, d\mathbf{W}_t \end{aligned}$$

Again, the equation preserves the length of the random process $\mathbf{X}_t$. It is significantly more complicated than equation (4.3). Indeed, two functions present in the equation, namely the potential V_e and the diffusion parameter B induce a substantial additional difficulty. The potential V_e models the effective interaction of the rod under consideration with the other rods, an interaction that is very important for liquid crystals. On the other hand, contrary to the situation of (4.3), the diffusion parameter B is not fixed, and the configuration dependence of the parameter B accounts for the fact that the rotational diffusivity of the rod depends on the orientation with respect to the other rods. The point is that both the potential V_e and the diffusion parameter B not only depend on $\mathbf{X}_t$, but depend on certain *averages* of $\mathbf{X}_t$ in the following way:

$$V_e(\mathbf{X}_t) = \mathrm{Cte} \quad \mathbb{E}\big(\big\|\mathbf{X}_t \times \mathbf{X}_t\big\|\big), \qquad (4.6)$$

$$B(\mathbf{X}_t) = \frac{\mathrm{Cte}}{\mathbb{E}\big(\big\|\mathbf{X}_t \times \mathbf{X}_t\big\|\big)}. \qquad (4.7)$$

This makes the stochastic differential equation (4.5) *nonlinear in the sense of MacKean* as its coefficients depend on averages of the solution itself. The setting considered in the present article does not cover this case.

5 The Fokker-Planck equation of divergence form

5.1 Some preliminaries

Let us begin with some notation. Fix a probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P}, \mathbf{W}_t)$ with a standard K-dimensional Brownian motion $\mathbf{W}_t$, a drift vector $\mathbf{b} \in \mathbb{R}^N$, a dispersion matrix $\sigma \in \mathbb{R}^{N\times K}$. To the generic stochastic differential equation

$$d\mathbf{X}_t = \mathbf{b}\,(\mathbf{X}_t)\,dt + \sigma(\mathbf{X}_t)\,d\mathbf{W}_t \tag{5.1}$$

that componentwise writes $d\mathbf{X}_t^i = \mathbf{b}_{\,i}(\mathbf{X}_t)\,dt + \sigma_{ij}(\mathbf{X}_t)\,d\mathbf{W}_t^j$, $1 \leq i \leq N$, we may first associate two partial differential equations. The Fokker-Planck (or forward Kolmogorov) equation

$$\partial_t\, p + \partial_i\,(pb_i) - \frac{1}{2}\,\partial^2_{ij}\,(\sigma_{ik}\sigma_{jk}\,p) = 0 \tag{5.2}$$

and its adjoint equation, the backward Kolmogorov equation

$$\partial_t\, f\; - b_i\partial_i\, f - \frac{1}{2}\,\sigma_{ik}\sigma_{jk}\partial^2_{ij}\, f = 0. \tag{5.3}$$

The parameter fields in the above equations are the drift vector $\mathbf{b}$ and the matrix σ. Throughout this section, and unless otherwise stated, they will be assumed to depend on space and not on time: $\mathbf{b} = \mathbf{b}\,(x)$, $\sigma = \sigma(x)$. Unless there is a risk of confusion, we will not explicitly mention their space dependence.

In the sequel (see Section 7), rather than the drift vector $\mathbf{b}$ itself, the following vector field

$$\mathbf{b}^\sigma\; = \mathbf{b}\; - \frac{1}{2}\,\mathrm{div}\,(\sigma\sigma^t) \tag{5.4}$$

i.e. in coordinates $b_i^\sigma = b_i - \frac{1}{2}\,\partial_j\left(\sigma_{ik}\sigma_{jk}\right)$, will play a crucial role for the study of equations (5.2) and (5.3). Its divergence reads

$$\mathrm{div}\,\mathbf{b}^\sigma\; = \partial_i\, b_i - \frac{1}{2}\,\partial^2_{ij}\left(\sigma_{ik}\sigma_{jk}\right) \tag{5.5}$$

Note that this modified drift vector $\mathbf{b}^\sigma$ is (in general) different from the drift vector that is obtained when considering the Stratonovich form of the stochastic differential equation (5.1), namely

$$d\mathbf{X}_t = \mathbf{b}^{\mathrm{Strat}}\,(\mathbf{X}_t)\,dt + \sigma(\mathbf{X}_t)\,\circ\,d\mathbf{W}_t \tag{5.6}$$

where

$$(\mathbf{b}^{\mathrm{Strat}}\,)_i = b_i - \frac{1}{2}\,\sigma_{jk}\partial_j\,\sigma_{ik} \tag{5.7}$$

The difference writes $(\mathbf{b}^\sigma\; - \mathbf{b}^{\mathrm{Strat}}\,)_i = -\frac{1}{2}\,\sigma_{ik}\,\partial_j\,\sigma_{jk}$.

In addition to (5.2) and (5.3), a third partial differential equation, formally connected to (5.1), may be considered:

$$\partial_t\, p + \partial_i\,(pb_i) - \frac{1}{2}\,\partial_i\,(\sigma_{ik}\sigma_{jk}\partial_j\; p) = 0. \tag{5.8}$$

This equation, that we will call here the *Fokker-Planck equation of divergence form*, is the easiest one to deal with (as compared to (5.2) and (5.3)) when it comes to existence and uniqueness issues. This owes to the fact that its second order differential operator $\partial_i\left(\sigma_{ik}\sigma_{jk}\partial_j\, p\right)$ has a self-adjoint form. The study of (5.8) is the purpose of the present section. We will indicate in Section 7 the necessary modifications of our arguments to address the same issues for (5.2) and (5.3) respectively. Note that the equation in divergence form (5.8) was extensively studied by H. Osada in [20], in a different perspective though. In the latter work, a general matrix a_{ij} replaces $\sigma_{ik}\sigma_{jk}$. Importantly, this matrix is assumed positive definite (but not necessarily symmetric). The a_{ij} are only assumed measurable. The coefficients b_i of the drift vector are taken of the form $b_i = \partial_i\, c_{ij}$, with all c_{ij} measurables. The consideration of the divergence form (5.8) is motivated in [20] by specific fluid mechanics applications (vortex processes,...).

Before we get to the heart of the matter, it is interesting to ask whether there is a connection, better than purely formal, between the partial differential equation (5.8) and the stochastic differential equation (5.1). The connection between the Fokker-Planck equation (5.2) and (5.1) is clear: by the Ito rule, the law of $\mathbf{X}_t$ solution to (5.1) is the solution to (5.2) starting from the law of $\mathbf{X}_t|_{t=0}$. That between the backward Kolmogorov equation (5.3) and (5.1) is also clear: the Feynman-Kac formula tells us that $\mathbb{E}(\varphi(\mathbf{X}_t^x))$ is the solution to (5.3) starting from φ when $\mathbf{X}_t^x$ is the solution to (5.1) starting from x. What about (5.8)? The only argument we are aware of in this direction is the following one, specific to the one-dimensional setting.

We begin by noticing that, for y fixed in $\mathbb{R}$, and $\mathbf{X}_t^y$ the solution to (5.1) starting from y at initial time, $p(t,x;y) = \mathbb{E}(\delta(x - \mathbf{X}_t^y))$ is the solution to (5.2) supplied with the initial condition $p(t=0,x;y) = \delta(y-x)$.

Recall the proof of the elementary fact. By the Feynman-Kac formula, we know that, for x fixed, the solution $q(t,y;x)$ to

$$\begin{cases} \partial_t\, q - b\,\partial_y q - \frac{1}{2}\sigma^2\partial_y q = 0 \\ q(t=0,y;x) = \delta(x-y) \end{cases}$$

(i.e. (5.3) with initial condition $\delta(x-\cdot)$) is $q(t,y;x) = \mathbb{E}(\delta(x - \mathbf{X}_t^y))$ where $\mathbf{X}_t^y$ the solution to (5.1) starting from y at initial time. On the other hand, u being fixed, the solution $p(t,z;u)$ to

$$\begin{cases} \partial_t\, p + \partial_z(bp) - \frac{1}{2}\partial_{zz}(\sigma^2 p) = 0 \\ p(t=0,z;u) = \delta(u-z) \end{cases}$$

(i.e. (5.2) with initial condition $\delta(u-\cdot)$) is the law of $\mathbf{X}_t^u$ (solution to (5.1) starting from u at initial time). Thus,

$$\int p(t,z;u)\varphi(z)\,dz = \mathbb{E}(\varphi(\mathbf{X}_t^u)),$$

for all regular φ. Applying this (formally) to $\varphi = \delta(x - \cdot)$ and $u = y$, we obtain

$$p(t,x;y) = \int q(t,z;y)\delta(x-z)\,dy = \mathbb{E}(\delta(x - \mathbf{X}_t^y)) = q(t,y;x).$$

Thus $p(t,x;y) = \mathbb{E}(\delta(x - \mathbf{X}_t^y))$ is the solution to

$$\begin{cases} \partial_t p + \partial_x(bp) - \frac{1}{2}\partial_{xx}(\sigma^2 p) = 0 \\ p(t=0,x;y) = \delta(y-x) \end{cases}$$

which is our claim.

We now specifically use the one dimensional setting and introduce

$$F(t,x;y) = \int_{-\infty}^{x} p(t,z;y)\,dz,$$

which is in fact $F(t,x;y) = \mathbb{P}(\mathbf{X}_t^y < x)$, because $p(t,z;y) = \mathbb{E}(\delta(z - \mathbf{X}_t^y))$. The equation satisfied by F is obtained by integrating (5.2) from $-\infty$ to x:

$$\partial_t F(t,x;y) + b(x)\partial_x F(t,x;y) - \frac{1}{2}\partial_x(\sigma^2 \partial_x F(t,x;y)) = 0$$

that is (5.8), with the initial condition $F(0,x;y) = \int_{-\infty}^{x} p(0,z;y)\,dz = 1_{y<x}$. We have thus shown that $\mathbb{P}(\mathbf{X}_t^y < x)$ is the solution to (5.8) starting from $1_{y<x}$. This establishes a natural connection between $\mathbf{X}_t$ and (5.8).

5.2 A priori estimate

Following the strategy outlined in the introduction, we begin by showing on (5.8) some formal *a priori* estimates. We formally multiply (5.8)

$$\partial_t p + \partial_i (pb_i) - \frac{1}{2}\partial_i (\sigma_{ik}\sigma_{jk}\partial_j\ p) = 0$$

by p and integrate on the whole space to get:

$$\frac{d}{dt}\int \frac{p^2}{2} + \int p\partial_i (pb_i) - \int \frac{1}{2}\partial_i (\sigma_{ik}\sigma_{jk}\partial_j\ p)\,p = 0.$$

Integrating by parts the last two terms, we obtain

$$\frac{d}{dt}\int \frac{p^2}{2} + \int \frac{p^2}{2}\operatorname{div}\mathbf{b} + \frac{1}{2}\int \left|\sigma^t(x)\nabla p\right|^2 = 0. \tag{5.9}$$

This estimate is somehow the key estimate that will play a central role throughout the article. Assuming

$$\operatorname{div}\mathbf{b} \in L^\infty \tag{5.10}$$

we obtain $p \in L^\infty([0,T], L^2 \cap L^\infty)$, $\sigma^t\nabla p \in L^2([0,T], L^2)$. The same assumption (5.10) also allows to get a L^∞ bound on p using the maximum principle. Therefore the solution p is expected to belong to the space

$$X_2 = \left\{p \in L^\infty([0,T], L^2 \cap L^\infty), \quad \sigma^t\nabla p \in L^2([0,T], L^2)\right\}. \tag{5.11}$$

This holds of course provided the initial condition lies in $L^2 \cap L^\infty$.

Remark 3 *In fact, we may remark that the condition*

$$\inf \operatorname{div} \mathbf{b} > -C \tag{5.12}$$

is sufficient to argue on (5.9). This is true throughout this article.

Remark 4 *We may also multiply (5.8) by $\beta'(p)$ and obtain*

$$\frac{d}{dt}\int \beta(p) + \int \big(p\beta'(p) - \beta(p)\big) \operatorname{div} \mathbf{b} + \frac{1}{2}\int \beta''(p)\big|\sigma^t(x)\nabla p\big|^2 = 0. \tag{5.13}$$

This may lead to a L^p theory, instead of the above L^2 theory. See Section 6 for further remarks.

5.3 Regularization

We now convoluate equation (5.8) with some regularizing kernel

$$\rho_\varepsilon = \varepsilon^{-N}\rho(\varepsilon^{-1}\cdot), \quad \text{with} \quad \rho \in \mathcal{D}(\mathbb{R}^N),\ \rho \geq 0,\ \int \rho = 1,$$

and obtain

$$\partial_t\, \rho_\varepsilon \star p + \rho_\varepsilon \star \partial_i\,(pb_i) - \frac{1}{2}\,\rho_\varepsilon \star \partial_i\,(\sigma_{ik}\sigma_{jk}\partial_j\; p) = 0. \tag{5.14}$$

Throughout this article, we will employ the notation (already introduced here in the previous section):

$$[\rho_\varepsilon, c](f) = \rho_\varepsilon \star (cf) - c\,(\rho_\varepsilon \star f),$$

for a differential operator c. We shall also denote $p_\varepsilon = \rho_\varepsilon \star p$. Using this notation, we thus have:

$$\begin{aligned}
\rho_\varepsilon \star \partial_i\,(pb_i) &= \rho_\varepsilon \star ((\partial_i\, b_i)p) + \rho_\varepsilon \star (b_i\partial_i\, p) \\
&= [\rho_\varepsilon, \partial_i\, b_i](p) + [\rho_\varepsilon, b_i\partial_i\,](p) + \partial_i\,(b_i p_\varepsilon) \\
&= Q_\varepsilon + \partial_i\,(b_i p_\varepsilon)
\end{aligned} \tag{5.15}$$

where we have defined

$$Q_\varepsilon = [\rho_\varepsilon, \partial_i\, b_i](p) + [\rho_\varepsilon, b_i\partial_i\,](p). \tag{5.16}$$

Likewise,

$$\begin{aligned}
\rho_\varepsilon \star \partial_i\,(\sigma_{ik}\sigma_{jk}\partial_j\; p) &= \partial_i\;(\rho_\varepsilon \star (\sigma_{ik}\sigma_{jk}\partial_j\; p)) \\
&= \partial_i\;([\rho_\varepsilon, \sigma_{ik}\sigma_{jk}\partial_j\,](p)) + \partial_i\;(\sigma_{ik}\sigma_{jk}\partial_j\; p_\varepsilon) \\
&= \partial_i\;(\sigma_{ik}[\rho_\varepsilon, \sigma_{jk}\partial_j\,](p) + [\rho_\varepsilon, \sigma_{ik}](\sigma_{jk}\partial_j\, p)) \\
&\qquad + \partial_i\;(\sigma_{ik}\sigma_{jk}\partial_j\; p_\varepsilon) \\
&= \partial_i\;(\sigma_{ik}[\rho_\varepsilon, \sigma_{jk}\partial_j\,](p)) + [\rho_\varepsilon, \partial_i\,\sigma_{ik}](\sigma_{jk}\partial_j\, p) \\
&\qquad + [\rho_\varepsilon, \sigma_{ik}\partial_i\,](\sigma_{jk}\partial_j\, p) + \partial_i\;(\sigma_{ik}\sigma_{jk}\partial_j\; p_\varepsilon) \\
&= \partial_i\;(\sigma_{ik}R_\varepsilon) + S_\varepsilon + T_\varepsilon + \partial_i\;(\sigma_{ik}\sigma_{jk}\partial_j\; p_\varepsilon)
\end{aligned} \tag{5.17}$$

where we have set

$$R_\varepsilon = [\rho_\varepsilon, \sigma_{jk}\partial_j\,](p), \quad S_\varepsilon = [\rho_\varepsilon, \partial_i\,\sigma_{ik}](\sigma_{jk}\partial_j\,p), \quad T_\varepsilon = [\rho_\varepsilon, \sigma_{ik}\partial_i\,](\sigma_{jk}\partial_j\,p) \tag{5.18}$$

We thus obtain from (5.14) the equation (5.8) set on p_ε but with an error term in the right-hand side:

$$\partial_t\,p_\varepsilon + \partial_i\,(p_\varepsilon b_i) - \frac{1}{2}\,\partial_i\,(\sigma_{ik}\sigma_{jk}\partial_j\;p_\varepsilon) = -Q_\varepsilon + \frac{1}{2}\,(\partial_i\;(\sigma_{ik}R_\varepsilon) + S_\varepsilon + T_\varepsilon)\,. \tag{5.19}$$

Remark 5 *Note that in the above manipulations we have essentially used the following two basic formulae, that can be readily proved:*

$$\partial\left([\rho_\varepsilon, a](f)\right) = [\rho_\varepsilon, (\partial a)](f) + [\rho_\varepsilon, a\partial](f), \tag{5.20}$$

$$[\rho_\varepsilon, a\,b\,\partial](f) = a\,[\rho_\varepsilon, b\,\partial](f) + [\rho_\varepsilon, a](b\,\partial f), \tag{5.21}$$

for any f, and any two scalar fields a and b whenever all the above terms make sense. The generalization to vector-valued fields a and b is straightforward.

With a view to preparing the proof of the uniqueness of the solution to (5.8), we now examine the right-hand side of (5.19). As in [17], our main tool is the following commutation lemma:

Lemma 1 (see [6, Lemma II.1] and [17, Lemma 5.1])

Let $f \in L^r_{loc}(\mathbb{R}^N)$ and let $\mathbf{c} \in \left(W^{1,\alpha}_{loc}(\mathbb{R}^N)\right)^N$. Set $\frac{1}{\beta} = \frac{1}{r} + \frac{1}{\alpha}$. Then, as $\varepsilon \longrightarrow 0$,

$$[\rho_\varepsilon, \mathbf{c}\cdot\nabla](f) \longrightarrow 0 \quad \text{in} \quad L^\beta_{loc}(\mathbb{R}^N), \tag{5.22}$$

and

$$[\rho_\varepsilon, \operatorname{div}\mathbf{c}\,](f) \longrightarrow 0 \quad \text{in} \quad L^\beta_{loc}(\mathbb{R}^N). \tag{5.23}$$

If $\nabla f \in L^2_{loc}(\mathbb{R}^N)$ and $\mathbf{c} \in \left(L^\alpha_{loc}(\mathbb{R}^N)\right)^N$, then the same conclusion holds for $\frac{1}{\beta} = \frac{1}{2} + \frac{1}{\alpha}$.

Remark 6 *In fact (5.23) only requires an assumption slightly weaker than $\mathbf{c} \in \left(W^{1,\alpha}_{loc}(\mathbb{R}^N)\right)^N$, namely $\operatorname{div}\mathbf{c} \in \left(L^\alpha_{loc}(\mathbb{R}^N)\right)^N$.*

Using the above lemma, we immediately see that, when $\mathbf{b} \in W^{1,1}$ and $\operatorname{div}\mathbf{b} \in L^\infty$, we may claim the convergence to zero of the first order error term

$$Q_\varepsilon = [\rho_\varepsilon, \partial_i\,b_i](p) + [\rho_\varepsilon, b_i\partial_i\,](p) \stackrel{\varepsilon\to 0}{\longrightarrow} 0, \qquad \text{in} \quad L^1_{loc} \tag{5.24}$$

This convergence holds uniformly in time, because p is L^∞ and $\mathbf{b}$ is time independent. This term was the standard term, and we now get to the, non standard, second order term. Applying again the above lemma, we have:

$$R_\varepsilon = [\rho_\varepsilon, \sigma_{jk}\partial_j\,](p) \qquad \stackrel{\varepsilon\to 0}{\longrightarrow} 0, \qquad \text{in} \quad L^2_{loc}, \tag{5.25}$$

as soon as $\sigma \in \left(W^{1,2}_{loc}\right)^{N\times K}$. Again, this convergence holds uniformly in time, for $p \in L^\infty([0,T],L^2)$. Likewise,

$$S_\varepsilon = [\rho_\varepsilon, \partial_i\, \sigma_{ik}](\sigma_{jk}\partial_j\, p) \qquad \stackrel{\varepsilon\to 0}{\longrightarrow} 0, \qquad \text{in} \quad L^1_{loc}, \tag{5.26}$$

when σ is again $W^{1,2}$ and $\sigma^t\nabla p \in \left(L^2\right)^N$. With respect to time, the convergence (5.26) holds in L^2, because $\sigma^t\nabla p \in L^2([0,T],L^2)$. Regarding the last term, we have

$$T_\varepsilon = [\rho_\varepsilon, \sigma_{ik}\partial_i\,](\sigma_{jk}\partial_j\, p) \qquad \stackrel{\varepsilon\to 0}{\longrightarrow} 0, \qquad \text{in} \quad L^1_{loc}, \tag{5.27}$$

also in L^2 with respect to time for the same reason as regards (5.26). Setting

$$U_\varepsilon = -Q_\varepsilon + \frac{1}{2}S_\varepsilon + \frac{1}{2}T_\varepsilon,$$

we now collect all these convergences in (5.19),

$$\begin{aligned} &\partial_t\, p_\varepsilon + \partial_i\,(p_\varepsilon b_i) - \frac{1}{2}\,\partial_i\,(\sigma_{ik}\sigma_{jk}\partial_j\; p_\varepsilon) = U_\varepsilon + \frac{1}{2}\partial_i\,\left(\sigma_{ik}R_\varepsilon\right), \\ &\quad\text{with} \quad U_\varepsilon \stackrel{\varepsilon\to 0}{\longrightarrow} 0 \quad \text{in} \quad L^\infty + L^2([0,T],L^1_{loc}) \\ &\quad\text{and} \quad R_\varepsilon \stackrel{\varepsilon\to 0}{\longrightarrow} 0 \quad \text{in} \quad L^\infty([0,T],L^2_{loc}), \end{aligned} \tag{5.28}$$

under the conditions

$$\mathbf{b} \in W^{1,1}_{loc}, \qquad \operatorname{div}\mathbf{b} \in L^\infty, \qquad \sigma \in \left(W^{1,2}_{loc}\right)^{N\times K}, \qquad \sigma^t\nabla p \in \left(L^2\right)^N. \tag{5.29}$$

We are now in position to prove our first existence and uniqueness result, on (5.8).

5.4 Existence and uniqueness

Based on the formal *a priori* estimate (5.9), we first define our notion of solution.

5.4.1 Setting of the equation

Let us be given an initial condition $p_0 \in L^2 \cap L^\infty(\mathbb{R}^N)$, we wish to build a solution p to the equation in $L^\infty([0,T],L^2\cap L^\infty(\mathbb{R}^N))$ satisfying in addition $\sigma^t\nabla p \in \left(\left(L^2([0,T],L^2(\mathbb{R}^N))\right)^N\right.$. Then the mathematical sense we give to equation (5.8) is the following

$$\begin{aligned} &-\int_0^T\int_{\mathbb{R}^N} p\,\frac{\partial\varphi}{\partial t} - \int_{\mathbb{R}^N} p_0\varphi(0,\cdot) - \int_0^T\int_{\mathbb{R}^N} p\,\mathbf{b}\,\cdot\nabla\varphi \\ &+\frac{1}{2}\int_0^T\int_{\mathbb{R}^N} \sigma^t\nabla p\,.\,\sigma^t(x)\nabla\varphi = 0, \end{aligned} \tag{5.30}$$

for all smooth φ with compact support in $[0,T[\times\mathbb{R}^N$. The above makes sense for

$$\mathbf{b} \in (L^1_{loc}(\mathbb{R}^N))^N, \qquad \sigma \in (L^2_{loc}(\mathbb{R}^N))^N. \tag{5.31}$$

5.4.2 Existence

Establishing existence of solution for (5.8) is, as announced in the introduction, the "easy" part. For this purpose, we first regularize the drift vector $\mathbf{b}$ and the dispersion matrix σ by considering $\mathbf{b}_\varepsilon = \rho_\varepsilon \star \mathbf{b}$, $\sigma_\varepsilon = \rho_\varepsilon \star \sigma$ and the equation

$$\partial_t p + \partial_i \left(p(b_\varepsilon)_i\right) - \frac{1}{2}\partial_i \left((\sigma_\varepsilon)_{ik}(\sigma_\varepsilon)_{jk}\partial_j \, p\right) = 0$$

for which, by standard arguments, a solution $p = p_\varepsilon$ exists. Using the *a priori* estimates, which are here not only formal but rigourous, we readily obtain that p_ε is a bounded sequence in X_2 (defined in (5.11)), provided we assume $\operatorname{div}\mathbf{b} \in L^\infty$. Extracting a subsequence, we pass to the limit and obtain a solution to (5.8), in the mathematical sense mentioned above.

5.4.3 Uniqueness

As regards uniqueness, we first argue, for pedagogic purposes, when the equation is posed on $[0,1]^N \subset \mathbb{R}^N$, with periodic boundary conditions. In this manner, we do not have to worry about boundary terms in integration by parts, and all local regularities and integrabilities are immediately global. We will explain, in a second stage, how our arguments have to be modified to treat the case of the whole space $\mathbb{R}^N$.

By linearity of the equation, we know that proving uniqueness amounts to proving that a solution p to (5.8) starting from zero at initial time, and belonging to $\left\{p \in L^\infty([0,T], L^2 \cap L^\infty), \quad \sigma^t \nabla p \in L^2([0,T], L^2)\right\}$, necessarily vanishes for all times.

Convoluating (5.8) by ρ_ε as in the previous section, we have

$$\partial_t \, p_\varepsilon + \partial_i \left(p_\varepsilon b_i\right) - \frac{1}{2}\partial_i \left(\sigma_{ik}\sigma_{jk}\partial_j \, p_\varepsilon\right) = U_\varepsilon + \frac{1}{2}\partial_i \, \left(\sigma_{ik}R_\varepsilon\right),$$

where $p_\varepsilon = \rho_\varepsilon \star p$ and the properties of U_ε and R_ε are defined in (5.28). Multiplying by p_ε and integrating by part on the unit cube, we obtain:

$$\begin{aligned} \frac{d}{dt}\int_{[0,1]^N} \frac{p_\varepsilon^2}{2} + \frac{1}{2}\int_{[0,1]^N} p_\varepsilon^2 \operatorname{div}\mathbf{b} + \frac{1}{2}\int_{[0,1]^N} \left|\sigma^t(x)\nabla p_\varepsilon\right|^2 \\ = \int_{[0,1]^N} U_\varepsilon \, p_\varepsilon - \frac{1}{2}\int_{[0,1]^N} \left(\sigma^t(x)\nabla p_\varepsilon\right) . \, R_\varepsilon \end{aligned} \tag{5.32}$$

Recalling that the initial value is zero, we integrate in time from 0 to t and obtain (in fact this is also a direct consequence of our notion of solutions):

$$\begin{aligned} \int_{[0,1]^N} \frac{p_\varepsilon^2}{2}(t) + \frac{1}{2}\int_0^t \int_{[0,1]^N} p_\varepsilon^2 \operatorname{div}\mathbf{b} + \frac{1}{2}\int_0^t \int_{[0,1]^N} \left|\sigma^t(x)\nabla p_\varepsilon\right|^2 \\ = \int_0^t \int_{[0,1]^N} U_\varepsilon \, p_\varepsilon - \frac{1}{2}\int_0^t \int_{[0,1]^N} \left(\sigma^t(x)\nabla p_\varepsilon\right) . \, R_\varepsilon \end{aligned} \tag{5.33}$$

We have

$$\left| \int_{\mathbf{R}^N} p_\varepsilon^2 \operatorname{div} \mathbf{b} \right| \leq \|\operatorname{div} \mathbf{b}\|_{L^\infty(\mathbf{R}^N)} \int_{\mathbf{R}^N} p_\varepsilon^2. \tag{5.34}$$

and of course

$$\int_0^t \int_{[0,1]^N} \left|\sigma^t(x)\nabla p_\varepsilon\right|^2 \geq 0 \tag{5.35}$$

On the other hand, since p_ε is bounded in L^∞ and $\sigma^t \nabla p_\varepsilon$ is bounded in L^2, the right-hand side goes to zero with ε in view of the convergences (5.28). Letting ε go to zero, we thus obtain:

$$\int_{[0,1]^N} \frac{p^2}{2} \leq C \int_0^t \int_{[0,1]^N} p^2,$$

and we finally get $p \equiv 0$, which concludes the proof.

To address the case of the whole space $\mathbb{R}^N$, we have to specify some growth conditions at infinity for both $\mathbf{b}$ and σ, in addition to the assumptions (5.29) which were sufficient for the periodic case. These growth conditions will appear in our argument below.

We first introduce a nonnegative smooth cut-off function ϕ_R such that $\phi_R = \phi(\frac{\cdot}{R})$, $\phi \equiv 1$ on the ball of radius R, ϕ vanishes outside the ball of radius 2. Note we then have $\nabla \phi_R = \frac{1}{R} \nabla \phi(\frac{\cdot}{R})$. Next, we essentially redo the above proof, multiplying this time the equation by $p_\varepsilon \phi_R$ and integrating over the space $\mathbb{R}^N$. Using the same manipulations as above, but this time keeping track of all boundary terms, we obtain

$$\begin{aligned}
&\frac{d}{dt} \int_{\mathbf{R}^N} \frac{p_\varepsilon^2}{2} \phi_R + \frac{1}{2} \int_{\mathbf{R}^N} p_\varepsilon^2 \phi_R \operatorname{div} \mathbf{b} + \frac{1}{2} \int_{\mathbf{R}^N} \left|\sigma^t(x)\nabla p_\varepsilon\right|^2 \phi_R \\
&\quad = \frac{1}{2} \int_{\mathbf{R}^N} p_\varepsilon^2 \, \mathbf{b} \, \nabla \phi_R - \frac{1}{2} \int_{\mathbf{R}^N} p_\varepsilon \, \sigma^t(x) \nabla p_\varepsilon \cdot \sigma^t(x) \nabla \phi_R \\
&\qquad + \int_{[0,1]^N} U_\varepsilon \, p_\varepsilon \phi_R - \frac{1}{2} \int_{\mathbf{R}^N} (\sigma^t(x)\nabla p_\varepsilon) \,.\, R_\varepsilon \phi_R \\
&\qquad - \frac{1}{2} \int_{\mathbf{R}^N} p_\varepsilon (\sigma^t(x) \nabla \phi_R) \,.\, R_\varepsilon
\end{aligned} \tag{5.36}$$

Of course, the terms with ϕ_R will in the end give the main contributions. They are treated as in the periodic case above. We therefore focus on the terms in $\nabla \phi_R$. First, we note

$$\begin{aligned}
\left| \frac{1}{2} \int_{\mathbf{R}^N} p_\varepsilon^2 \, \mathbf{b} \, \nabla \phi_R \right| &= \left| \frac{1}{2} \, \frac{1}{R} \int_{R \leq |x| \leq 2R} p_\varepsilon^2 \, \mathbf{b} \, (\nabla \phi)(\frac{\cdot}{R}) \right| \\
&\leq C \left\| \frac{\mathbf{b}}{1+|x|} \right\|_{L^\infty} \|\nabla \phi\|_{L^\infty} \int_{|x| \geq R} p_\varepsilon^2.
\end{aligned} \tag{5.37}$$

This goes to zero as $R \longrightarrow +\infty$, uniformly in $\varepsilon \leq 1$, and L^∞ in time. Likewise,

$$\left| \int_{\mathbf{R}^N} p_\varepsilon \, \sigma^t(x) \nabla p_\varepsilon \cdot \sigma^t(x) \nabla \phi_R \right| \leq C \left\| \frac{\sigma}{1+|x|} \right\|_{L^\infty} \|\nabla \phi\|_{L^\infty} \int_{|x| \geq R} |p_\varepsilon| \, |\sigma^t \nabla p_\varepsilon| \tag{5.38}$$

also goes to zero as $R \longrightarrow +\infty$, uniformly in $\varepsilon \leq 1$.

These two terms can be collected into a term $W(t, \varepsilon, R)$ that goes to zero as $R \longrightarrow +\infty$, uniformly in $\varepsilon \leq 1$ and L^1 in time (because of (5.28)).

Regarding the last term of (5.36), we denote it by $V(t, \varepsilon, R)$ and bound it as follows

$$\begin{aligned} |V(t,\varepsilon,R)| &= \left| \int_{\mathbf{R}^N} p_\varepsilon (\sigma^t(x) \nabla \phi_R) \, . \, R_\varepsilon \right| \\ &\leq C \left\| \frac{\sigma}{1+|x|} \right\|_{L^\infty} \|\nabla \phi\|_{L^\infty} \int_{|x| \leq 2R} |p_\varepsilon| \, |R_\varepsilon| \end{aligned} \tag{5.39}$$

We thus recover (5.32) up to this remainder term $V(t, \varepsilon, R) + W(t, \varepsilon, R)$, and thus, instead of (5.33):

$$\begin{aligned} \int_{[0,1]^N} \frac{p_\varepsilon^2}{2} \phi_R(t) \quad &+ \quad \frac{1}{2} \int_0^t \int_{[0,1]^N} p_\varepsilon^2 \mathrm{div}\, \mathbf{b} \; \phi_R + \frac{1}{2} \int_0^t \int_{[0,1]^N} \left| \sigma^t(x) \nabla p_\varepsilon \right|^2 \phi_R \\ &= \int_0^t \int_{[0,1]^N} U_\varepsilon \, p_\varepsilon - \frac{1}{2} \int_0^t \int_{[0,1]^N} (\sigma^t(x) \nabla p_\varepsilon) \, . \, R_\varepsilon \\ &\quad + \int_0^t V(s, \varepsilon, R) \, ds \\ &\quad + \int_0^t W(s, \varepsilon, R) \, ds \end{aligned} \tag{5.40}$$

We then fix $\eta > 0$. We may find R large enough so that, uniformly in $\varepsilon \leq 1$, we have

$$\left| \int_0^t W(s, \varepsilon, R) ds \right| \leq \eta$$

for all $t \leq T$. For such a radius R, we let ε go to zero. The first three terms of the right-hand side of (5.40) then vanish, owing to (5.39) and the convergences (5.28) established above. In the limit we thus obtain

$$\int_{\mathbf{R}^N} \frac{p^2}{2} \, \phi_R \leq \eta + C \int_0^t \int_{\mathbf{R}^N} p^2 \, \phi_R$$

thus

$$\int_0^t \int_{\mathbf{R}^N} p^2 \, \phi_R \leq \frac{\eta}{C} (e^{Ct} - 1).$$

We now let R go to infinity and obtain

$$\int_0^t \int_{\mathbf{R}^N} p^2 \leq \frac{\eta}{C} (e^{Ct} - 1).$$

It remains to let η go to zero, and we finally get $p \equiv 0$, which concludes the proof of the following proposition.

Proposition 1 [Fokker-Planck equation of divergence form] *Assume* $\mathbf{b}$ *and* σ *are time-independent and satisfy:*

$$\mathbf{b} \in \left(W^{1,1}_{loc}\right)^N, \quad \operatorname{div} \mathbf{b} \in L^\infty, \quad \frac{\mathbf{b}}{1+|x|} \in (L^\infty)^N \tag{5.41}$$

$$\sigma \in \left(W^{1,2}_{loc}\right)^{N\times K}, \quad \frac{\sigma}{1+|x|} \in (L^\infty)^{N\times K} \tag{5.42}$$

Then, for each initial condition in $L^2 \cap L^\infty$, equation (5.8) has a unique solution in the space

$$\left\{p \in L^\infty([0,T], L^2 \cap L^\infty), \sigma^t \nabla p \in L^2([0,T], L^2)\right\}.$$

Remark 7 *It should be noticed that the above conditions on the growth at infinity of* $\mathbf{b}$ *and* σ *contained in (5.41)-(5.42) agree with the conditions (3.7)-(3.8) of Theorem 2 (recalled in Section 3.2) on the martingale problem. They also agree with the conditions (3.4)-(3.5) of Theorem 1 on pathwise existence and uniqueness (recalled in Section 3.3).*

Remark 8 *In Section 8.1 below, we shall comment upon the optimality of the $W^{1,2}$ regularity of the dispersion matrix σ.*

Remark 9 *A consequence of the above theorem is a stability result for solutions to (5.8). We skip the statement of this stability result, which is similar to those of [6, Theorem II.4] and [17, Theorem 3.1].*

6 Variants and extensions

As announced above, we collect in this section several variants and extensions of the previous arguments, which allow to prove existence and uniqueness results similar to Proposition 1, all on equation (5.8). On the other hand, Section 7 will present the application of our arguments to equations different from (5.8).

6.1 Immediate extensions

6.1.1 Growth at infinity

We first would like to extend our result to the case when the growth at infinity of $\mathbf{b}$ and σ is somewhat more general that the $(1+|x|)L^\infty$ behavior assumed in (5.41)-(5.42).

For this purpose we return to (5.37)-(5.38)-(5.39), successively. We may write, instead of (5.37)

$$\begin{aligned}
\left|\int_{\mathbf{R}^N} p_\varepsilon^2 \, \mathbf{b} \, \nabla\phi_R\right| &= \left|\frac{1}{R}\int_{R\le|x|\le 2R} p_\varepsilon^2 \, \mathbf{b} \, (\nabla\phi)\left(\frac{\cdot}{R}\right)\right| \\
&\le C \, \|p_\varepsilon\|^2_{L^\infty} \, \|\nabla\phi\|_{L^\infty} \int_{|x|\ge R}\left|\frac{\mathbf{b}}{1+|x|}\right|,
\end{aligned} \tag{6.1}$$

whenever this last term makes sense and conclude our argument of Section 5.4 without any further change (Note that the right-hand side of (6.1), as that of (5.37), converges to zero L^∞ in time). By linearity we may thus extend Proposition 1 to the case

$$\frac{\mathbf{b}}{1+|x|} \in L^1 + L^\infty \tag{6.2}$$

Likewise, (5.38) may be replaced by

$$\begin{aligned} & \left| \int_{\mathbf{R}^N} p_\varepsilon \, \sigma^t(x) \nabla p_\varepsilon \cdot \sigma^t(x) \nabla \phi_R \right| \\ \leq \; & C \, \|p_\varepsilon\|_{L^\infty} \, \|\nabla \phi\|_{L^\infty} \int_{|x| \geq R} \left| \frac{\sigma}{1+|x|} \right| \, |\sigma^t \nabla p_\varepsilon| \\ \leq \; & C \, \|p_\varepsilon\|_{L^\infty} \, \|\nabla \phi\|_{L^\infty} \, \left\|\sigma^t \nabla p_\varepsilon\right\|_{L^2} \left(\int_{|x| \geq R} \left| \frac{\sigma}{1+|x|} \right|^2 \right)^{1/2}, \end{aligned} \tag{6.3}$$

when we assume $\dfrac{\sigma}{1+|x|} \in L^2$. Again, the right-hand side of (6.3) goes to zero as $R \longrightarrow +\infty$, in the L^1 sense in time, exactly as that of (5.38).

Under the same assumption on σ, we also have, instead of (5.39)

$$\begin{aligned} |V(t, \varepsilon, R)| &= \left| \int_{\mathbf{R}^N} p_\varepsilon (\sigma^t(x) \nabla \phi_R) \,.\, R_\varepsilon \right| \\ &\leq C \, \|p_\varepsilon\|_{L^\infty} \, \|\nabla \phi\|_{L^\infty} \int_{|x| \leq 2R} \left| \frac{\sigma}{1+|x|} \right| \, |R_\varepsilon| \end{aligned} \tag{6.4}$$

and the argument proceeds further (for R fixed sufficiently large, the term vanishes in the limit $\varepsilon \longrightarrow 0$). Therefore

$$\frac{\sigma}{1+|x|} \in L^2 + L^\infty \tag{6.5}$$

also allows to conclude.

Remark 10 *Note that the assumption* $\dfrac{\sigma}{1+|x|} \in L^1$ *does not* a priori *allow to conclude. It seems impossible, in this case, to control the term* $\int_{\mathbf{R}^N} p_\varepsilon \, \sigma^t(x) \nabla p_\varepsilon \cdot \sigma^t(x) \nabla \phi_R$ *in a manner similar to (5.38) or (6.3). On the other hand, the term* $V(t, \varepsilon, R)$ *(which is in fact treated as a term at finite distance) may be easily handled. This should not be a surprising fact. Assumption (6.5) is the* natural *assumption, since it exactly means* $\dfrac{a}{1+|x|^2} \in L^1 + L^\infty$.

6.1.2 L^1 theory

In all Section 5 we have assumed that the initial condition p_0 is $L^2 \cap L^\infty$ and have thus constructed a L^2 theory, as exemplified by the *a priori* estimate (5.9).

Let us indicate here the modification of our argument needed to address the case

$$p_0 \in L^1 \cap L^\infty. \tag{6.6}$$

To obtain the propagation of the L^1 regularity, we first make use of the *a priori* estimate (5.13) for a sequence of convex regular functions β converging to the absolute value function, and deduce from

$$\frac{d}{dt}\int \beta(p) + \int \big(p\beta'(p) - \beta(p)\big)\,\mathrm{div}\,\mathbf{b} \le 0,$$

that

$$\frac{d}{dt}\int |p| \le \int |p|\,|\mathrm{div}\,\mathbf{b}\,|\,. \tag{6.7}$$

This, (5.9) and the L^∞ bound on p obtained from the maximum principle, show that

$$p \in X_1 := \left\{p \in L^\infty([0,T], L^1 \cap L^\infty), \quad \sigma^t \nabla p \in L^2([0,T], L^2)\right\}, \tag{6.8}$$

for all times.

Since it only makes use of bounds on p that are local in space, the regularization argument of Section 5 is not modified. For the same reason, the existence proof is not modified either, apart from (of course) the functional space that is considered (X_1 instead of X_2). The notion of solution is modified correspondingly. As regards uniqueness, it can be verified that the *global* L^1 integrability of p does not allow for any (even slight) simplification of the argument made, nor any improvement of the assumptions on $\mathbf{b}$ and σ.

6.1.3 Time-dependent b and σ

We mentioned very early in the present article that the case of time-independent drift vectors $\mathbf{b}$ and dispersion matrix σ was considered for simplicity. As was the case in [6, 17] and many other works, time-dependent coefficients may be dealt with, as soon as all the regularity and integrability assumptions we made are taken in the L^1 sense in time. More precisely, it is well known, and first proved in [6], that we may assume

$$\begin{aligned} &\mathbf{b} \in \Big(L^1([0,T], W^{1,1}_{loc}(\mathbb{R}^N))\Big)^N, \quad \mathrm{div}\,\mathbf{b} \in L^1([0,T], L^\infty(\mathbb{R}^N)),\\ &\frac{\mathbf{b}}{1+|x|} \in \big(L^1([0,T], L^1 + L^\infty(\mathbb{R}^N))\big)^N \end{aligned} \tag{6.9}$$

(instead of (5.41) for time-independent drift vectors) and all the results on transport equations are preserved, up to slight modifications of the proofs. The crucial step is the consideration of the error term Q_ε defined in (5.16). It is easily shown that a L^1 dependence of $\|\mathrm{div}\,\mathbf{b}\,\|_{L^\infty}$ and $\|\nabla \mathbf{b}\,\|_{L^1_{loc}}$ allows to conclude that Q_ε vanishes in $L^1([0,T], L^1_{loc})$ as ε goes to zero. We have shown in [17] that the same setting is also convenient for (5.8) when $\sigma \equiv \mathrm{Id}$. And since the above proof of Proposition 1 mimicks the arguments of [6, 17], this again holds true here (Note that all terms in $\mathbf{b}$ and σ are treated separately in the above argument).

The slight novelty regards the time-dependence of σ. We claim that the assumption

$$\sigma \in \Big(L^2([0,T], W^{1,2}_{loc})\Big)^{N\times K}, \quad \frac{\sigma}{1+|x|} \in \big(L^2([0,T], L^2 + L^\infty(\mathbb{R}^N))\big)^{N\times K} \tag{6.10}$$

allows to conclude to existence and uniqueness. Let us make this precise.

It is immediate to note that the time-dependence of σ does not modify at all the *a priori* estimate. In the regularization step, we only have to consider the error terms R_ε, $S_\varepsilon =$, T_ε. Let us begin with the last two, $S_\varepsilon = [\rho_\varepsilon, \partial_i\, \sigma_{ik}](\sigma_{jk}\partial_j\, p)$ and $T_\varepsilon = [\rho_\varepsilon, \sigma_{ik}\partial_i\,](\sigma_{jk}\partial_j\, p)$. The least we can ask is that they vanish in $L^1([0,T], L^1_{loc}(\mathbb{R}^N))$. For this to hold true, since we only know that $\sigma^t\nabla p$ is $L^2([0,T], L^2(\mathbb{R}^N))$, we need that the $W^{1,2}$ regularity of σ holds in the L^2 sense in time, that is $\sigma \in \Big(L^2([0,T], W^{1,2}_{loc})\Big)^{N\times K}$. The same assumption allows to treat the term $R_\varepsilon = [\rho_\varepsilon, \sigma_{jk}\partial_j\,](p)$, that vanishes in $L^2([0,T], L^2_{loc}(\mathbb{R}^N))$ then, and thus can be conveniently integrated against $\sigma^t\nabla p_\varepsilon \in L^2([0,T], L^2(\mathbb{R}^N))^K$.

In the uniqueness proof, the above argument shows that all terms at finite distance may be treated. For the terms containing $\nabla\phi_R$, the same L^2 dependence in time of the assumption on $\dfrac{\sigma}{1+|x|}$ allow to conclude, as is easily seen when considering the right-hand sides of (5.38), (5.39), (6.3), (6.4).

All the above extensions are summarized in the following

Proposition 2 [FP equation of divergence form, general case] *Assume* $\mathbf{b}$ *and* σ *satisfy:*

$$\begin{aligned} &\mathbf{b} \in \Big(L^1([0,T], W^{1,1}_{loc}(\mathbb{R}^N))\Big)^N, \quad \operatorname{div}\mathbf{b} \in L^1([0,T], L^\infty(\mathbb{R}^N)),\\ &\frac{\mathbf{b}}{1+|x|} \in \big(L^1([0,T], L^1 + L^\infty(\mathbb{R}^N))\big)^N \end{aligned} \tag{6.11}$$

$$\sigma \in \Big(L^2([0,T], W^{1,2}_{loc}(\mathbb{R}^N))\Big)^{N\times K}, \frac{\sigma}{1+|x|} \in \big(L^2([0,T], L^2 + L^\infty(\mathbb{R}^N))\big)^{N\times K} \tag{6.12}$$

Then, for each initial condition in $L^1\cap L^\infty$ (resp. $L^2\cap L^\infty$), equation (5.8) has a unique solution in the space

$$\Big\{p \in L^\infty([0,T], L^1\cap L^\infty)(\text{resp.}\quad L^\infty([0,T], L^1\cap L^\infty)),\\ \sigma^t\nabla p \in L^2([0,T], L^2)\Big\}.$$

Remark 11 *Note that only $\sigma \in \Big(L^1([0,T], W^{1,2}_{loc})\Big)^{N\times K}$ does not allow to conclude. Again, as in Remark 10, this is a natural fact.*

Remark 12 *The generalization of the results of the present work to drift vectors* $\mathbf{b}$ *that are BV functions, in the spirit of the works [1, 2, 3, 16], is an interesting open question.*

6.2 A Girsanov transform in law

We now investigate an extension of Proposition 1 very different in spirit from the extensions considered above. The specific interest of this extension is that it has an intimate connection with some questions in the theory of stochastic differential equations.

We intend to prove:

Proposition 3 [Girsanov transform] *All the results of Proposition 2 hold true when* σ *satisfies (6.12) and*

$$\mathbf{b} = \widetilde{\mathbf{b}} + \sigma\,\theta, \quad \theta \in L^2([0,T], L^2 + L^\infty(\mathbb{R}^N))^K, \quad \widetilde{\mathbf{b}} \quad \text{satisfies} \quad (6.11). \tag{6.13}$$

The proof of Proposition 3 is rather straightforward. First we indicate how the proof of the *a priori* estimate is modified. We remark that

$$\begin{aligned}\int_0^T \int b_i \partial_i \frac{p^2}{2} &= \int_0^T \int \widetilde{b}_i \partial_i \frac{p^2}{2} + \int_0^T \int \sigma_{ik}\theta_k \,.\, \partial_i\, p\, p \\ &= \int_0^T \int \widetilde{b}_i \partial_i \frac{p^2}{2} + \int_0^T \int \sigma^t \nabla p \,.\, \theta p\end{aligned}$$

where the correction term is estimated using Hölder inequality:

$$\begin{aligned}\left|\int_0^T \int \sigma^t \nabla p \,.\, \theta p\right| &\le \int_0^T \left\|\sigma^t \nabla p\right\|_{L^2(\mathbb{R}^N)} \|\theta\|_{L^2+L^\infty(\mathbb{R}^N)} \|p\|_{L^\infty(\mathbb{R}^N)} \\ &\le \left\|\sigma^t \nabla p\right\|_{L^2([0,T],L^2(\mathbb{R}^N))} \|\theta\|_{L^2([0,T],L^2+L^\infty(\mathbb{R}^N))} \times \dots \\ &\qquad \|p\|_{L^\infty([0,T],L^2\cap L^\infty(\mathbb{R}^N))}\,.\end{aligned}$$

Consequently, this correction term does not modify the *a priori* estimate (5.8).

Second, the regularization procedure requests the consideration of the additional error term

$$\begin{aligned}\rho_\varepsilon \star \sigma_{ik}\theta_k \,.\, \partial_i\, p - \sigma_{ik}\theta_k \,.\, \partial_i\, p_\varepsilon &= [\rho_\varepsilon, \sigma_{ik}\theta_k \,.\, \partial_i\,](p), \\ &= \theta_k[\rho_\varepsilon, \sigma_{ik} \,.\, \partial_i\,](p) + [\rho_\varepsilon, \theta_k](\sigma_{ik}\partial_i\, p), \\ &= \theta \,.\, Y_\varepsilon + Z_\varepsilon, \end{aligned} \tag{6.14}$$

We have

$$Y_\varepsilon = [\rho_\varepsilon, \sigma^t \nabla](p) \overset{\varepsilon \longrightarrow 0}{\longrightarrow} 0, \quad \text{in} \quad L^2([0,T], L^2_{loc}(\mathbb{R}^N))^K \tag{6.15}$$

because $\sigma \in L^2([0,T], W^{1,2}_{loc}(\mathbb{R}^N))^{N\times K}$ and p is $L^\infty([0,T]\times\mathbb{R}^N)$. On the other hand,

$$Z_\varepsilon = [\rho_\varepsilon, \theta](\sigma^t \nabla p) \overset{\varepsilon \longrightarrow 0}{\longrightarrow} 0, \quad \text{in} \quad L^1([0,T], L^1_{loc}(\mathbb{R}^N)) \tag{6.16}$$

because $\sigma^t \nabla p \in L^2([0,T], L^2(\mathbb{R}^N))^K$ and $\theta \in L^2([0,T], L^2_{loc}(\mathbb{R}^N))^K$. These two convergences show that the error term (6.14) vanishes, as ε goes to zero, in $L^1([0,T], L^1_{loc}(\mathbb{R}^N))$. This allows to perform the regularization step as above.

Third, we have to check that the modification to (6.1) namely

$$\int_0^T \int_{\mathbb{R}^N} p_\varepsilon^2 \, \sigma \, \theta \, . \, \nabla \phi_R \leq C \, \|\nabla \phi\|_{L^\infty} \int_{|x| \geq R} p_\varepsilon^2 \left| \frac{\sigma}{1+|x|} \, \theta \right|, \tag{6.17}$$

vanishes for R large enough, uniformly in ε. This holds because $\nabla \phi$ is L^∞, and

$$\begin{aligned} p^2 \in L^\infty([0,T], L^1 \cap L^\infty(\mathbb{R}^N)), \\ \frac{\sigma}{1+|x|} \in \left(L^2([0,T], L^2 + L^\infty(\mathbb{R}^N))\right)^{N \times K}, \\ \theta \in L^2([0,T], L^2 + L^\infty(\mathbb{R}^N))^K. \end{aligned}$$

This concludes the proof of Proposition 3.

Remark 13 *Note that the assumption (6.2) on* $\mathbf{b}$ *, and the assumptions (6.5) on* σ *and (6.13) on* θ *are consistent with one another, since when comparing* $\sigma\theta$ *to* $\mathbf{b}$ *, we have:*

$$\frac{\sigma}{1+|x|} \in L^2 + L^\infty \quad \text{and} \quad \theta \in L^2 + L^\infty \quad \text{yields} \quad \frac{\sigma\theta}{1+|x|} \in L^1 + L^\infty,$$

which "agrees" with (6.2).

Remark 14 *When we wish to consider* $p \in L^1 + L^\infty$ *instead of* $p \in L^2 + L^\infty$*, as in Section 6.1.2, it is simple to deal with* $\theta \in L^2(\mathbb{R}^N)$*. Indeed,*

$$\left| \int_0^T \int \sigma_{ik} \theta_k \, . \, \partial_i \, p \right| \leq \left\| \sigma^t \nabla p \right\|_{L^2([0,T], L^2(\mathbb{R}^N))} \|\theta\|_{L^2([0,T], L^2 + L^\infty(\mathbb{R}^N))}, \tag{6.18}$$

and thus (6.7) is valid, up to a correction term that does not modify the argument. On the other hand, the adaptation is unclear for $\theta \in L^2 + L^\infty(\mathbb{R}^N)$.

Remark 15 *We do not know how to treat a case when* $\dfrac{\theta}{1+|x|} \in L^2 + L^\infty$.

It is now useful to explain how we may interpret, in terms of probability theory, the invariance of our assumptions on $\mathbf{b}$ with respect to the addition of $\sigma \, \theta$. For this purpose, we recall the well known Girsanov theorem. Consider $(\mathbf{X}_t, \mathbf{W}_t)$ a solution to the stochastic differential equation

$$d\mathbf{X}_t = (\mathbf{b} \, (\mathbf{X}_t) + \sigma \, \theta(\mathbf{X}_t)) \, dt + \sigma(\mathbf{X}_t) d\mathbf{W}_t, \tag{6.19}$$

for $\mathbf{b}$, σ regular and θ in L^2. Next we may set

$$\overline{\mathbf{W}_t} = \mathbf{W}_t + \int_0^t \theta(\mathbf{X}_s) \, ds, \tag{6.20}$$

which is a (K-dimensional) brownian motion under the probability $\mathbb{P}^\theta$ defined by

$$\frac{d\mathbb{P}^\theta}{d\mathbb{P}} = \exp\left(\int_0^t \theta(\mathbf{X}_s)\,.\,d\mathbf{W}s - \frac{1}{2}\int_0^t |\theta|^2(\mathbf{X}_s)\,ds\right) \tag{6.21}$$

Then, $(\mathbf{X}_t, \overline{\mathbf{W}_t})$ is a (weak) solution to the stochastic differential equation

$$d\mathbf{X}_t = \mathbf{b}\,(\mathbf{X}_t)\,dt + \sigma(\mathbf{X}_t)d\overline{\mathbf{W}_t}. \tag{6.22}$$

The Girsanov transform therefore amounts to replacing the drift vector $\mathbf{b}$ by $\mathbf{b} - \sigma\,\theta$ so that each given trajectory $\mathbf{X}_t$ still solves the equation (but for another brownian, now implicitly depending on $\mathbf{X}_t$).

In our context, the invariance with respect to the addition of $\sigma\,\theta$ could thus be seen as a Girsanov transform for *laws* rather than trajectories.

6.3 Full rank dispersion matrices

A specific argument may be done when the matrix $\sigma\sigma^t$ is uniformly positive definite i.e. there exists some constant $C_0 > 0$ such that

$$\left|\sigma^t(t,x)\zeta\right|^2 \geq C_0\left|\zeta\right|^2 \text{ for all}\,\zeta \in \mathbb{R}^N \text{ and for almost all}\,(t,x) \in \mathbb{R}_+ \times \mathbb{R}^N. \tag{6.23}$$

Let us as above begin with the formal *a priori* estimate. Inequality (6.23) clearly implies that

$$\int_{\mathbb{R}^N} \left|\sigma^t \nabla p\right|^2 \geq C_0 \int_{\mathbb{R}^N} \left|\nabla p\right|^2, \tag{6.24}$$

and thus, assuming $\operatorname{div}\mathbf{b} \in L^1([0,T], L^\infty(\mathbb{R}^N))$, the *a priori* estimate (5.9) implies that the natural functional space to consider is

$$X_{2,def} = \left\{p \in L^\infty([0,T], L^2 \cap L^\infty) \cap L^2([0,T], H^1),\, \sigma^t\nabla p \in L^2([0,T], L^2)\right\}, \tag{6.25}$$

instead of (5.11).

We now turn to the regularization step. With such a H^1 integrability on the solution, it is well known from [17] that the regularization step can then accomodate a drift vector that is only $L^2([0,T], L^2_{loc}(\mathbb{R}^N))$ and not necessarily $W^{1,1}$ in space (see the last assertion of Lemma 1). The error term for the first order operator is thus expressed as (5.15) and treated as in [17]. As regards the error term associated to the second order operator, we write, instead of (5.17),

$$\rho_\varepsilon \star \partial_i\left(\sigma_{ik}\sigma_{jk}\partial_j\; p\right) - \partial_i\left(\sigma_{ik}\sigma_{jk}\partial_j\; p_\varepsilon\right) = \partial_i\left([\rho_\varepsilon, \sigma_{ik}\sigma_{jk}\partial_j\,](p)\right) \tag{6.26}$$

where we now denote by $\partial_i\,R_\varepsilon$ the right-hand side. If we next assume

$$\sigma \in L^\infty([0,T], L^\infty_{loc}(\mathbb{R}^N)^{N\times K}$$

we see, by application of Lemma 1, that

$$R_\varepsilon = [\rho_\varepsilon, \sigma_{ik}\sigma_{jk}\partial_j\,](p) \overset{\varepsilon\longrightarrow 0}{\longrightarrow} 0, \quad \text{in} \quad L^2([0,T], L^2_{loc}(\mathbb{R}^N)) \tag{6.27}$$

Multiplying the regularized equation by $p_\varepsilon \phi_R$ and integrating in space and time yields the following specific contribution

$$\int_0^T \int [\rho_\varepsilon, \sigma_{ik}\sigma_{jk}\partial_j\,](p)\,\partial_i\,(p\phi_R) = \int R_\varepsilon(\phi_R\,\partial_i\,p + p\,\partial_i\,\phi_R) \tag{6.28}$$

where both terms of the right-hand side are shown to vanish when ε goes to zero (for R fixed) because

$$\phi_R \in W^{1,\infty},\ p \in L^2([0,T], H^1(\mathbb{R}^N)),\ R_\varepsilon \overset{\varepsilon \longrightarrow 0}{\longrightarrow} 0 \,\text{in}\, L^2([0,T], L^2_{loc}(\mathbb{R}^N)).$$

Consequently, the regularization may be performed, and the rest of the argument for existence and uniqueness follows unchanged. The following proposition is thus proved.

Proposition 4 [Positive definite diffusion matrices] *Assume that the matrix $\sigma\sigma^t$ is uniformly positive definite (i.e. that (6.23) holds) and that*

$$\begin{cases} \mathbf{b} \in L^2([0,T], L^2_{loc}(\mathbb{R}^N))^N, \qquad \dfrac{\mathbf{b}}{1+|x|} \in \left(L^1([0,T], L^1 + L^\infty(\mathbb{R}^N))\right)^N, \\ \operatorname{div}\mathbf{b} \in L^1([0,T], L^\infty(\mathbb{R}^N)) \\ \sigma \in L^\infty([0,T], L^\infty_{loc}(\mathbb{R}^N)^{N\times K}, \quad \dfrac{\sigma}{1+|x|} \in \left(L^2([0,T], L^2 + L^\infty(\mathbb{R}^N))\right)^{N\times K} \end{cases} \tag{6.29}$$

Then, for each initial condition in $L^1 \cap L^\infty$ (resp. $L^2 \cap L^\infty$), equation (5.8) has a unique solution in the space

$$\Big\{p \in L^\infty([0,T], L^2 \cap L^\infty)(\text{resp.} \quad L^\infty([0,T], L^1 \cap L^\infty)),$$
$$p \in L^2([0,T], H^1), \quad \sigma^t\nabla p \in L^2([0,T], L^2)\Big\}.$$

Remark 16 *Instead of the two assumptions* $\mathbf{b} \in L^2([0,T], L^2_{loc}(\mathbb{R}^N))^N$, $\operatorname{div}\mathbf{b} \in L^1([0,T], L^\infty(\mathbb{R}^N))$ *we may equivalently assume*

$$\mathbf{b} \in L^2([0,T], L^2(\mathbb{R}^N))^N,$$

(note the global L^2 *integrability). This owes to the majoration:*

$$\left|\int_0^T \int \mathbf{b}\,.\,\nabla f\, f\right| \le \|\mathbf{b}\|_{L^2([0,T],L^2(\mathbb{R}^N))}\, \|\nabla f\|_{L^2([0,T],L^2(\mathbb{R}^N))}\, \|f\|_{L^\infty([0,T]\times\mathbb{R}^N))}, \tag{6.30}$$

which then allows for establishing the a priori *estimate. On the other hand, the regularization step is not modified since it consists of a local argument.*

Remark 17 *All the extensions of the results of the previous section may be applied* mutatis mutandis *to the above proposition. Also, by linearity, we may consider a drift vector that is a sum of $\mathbf{b}^1$ satisfying the assumptions of Proposition 4 and $\mathbf{b}^2$ satisfying those of Proposition 2 (or of its extension Proposition 3).*

7 Adaptation to Fokker-Planck type equations

We devote this section to the adaptation of the above results (on (5.8)) to the Fokker-Planck equation (5.2) and the backward Kolmogorov equation (5.3). Actually, it will be seen that the adaptation is straightforward, up to a modification of the drift vector $\mathbf{b}$ into the drift vector $\mathbf{b}^\sigma$ defined in (5.4).

Considering that

$$-\frac{1}{2}\,\partial^2_{ij}\left(\sigma_{ik}\sigma_{jk}\,p\right) = -\frac{1}{2}\,\partial_i\left(\sigma_{ik}\sigma_{jk}\partial_j\ p\right) - \frac{1}{2}\,\partial_i\left(\partial_j\left(\sigma_{ik}\sigma_{jk}\right)p\right) \tag{7.1}$$

the Fokker-Planck equation (5.2)

$$\partial_t\,p + \partial_i\left(pb_i\right) - \frac{1}{2}\,\partial^2_{ij}\left(\sigma_{ik}\sigma_{jk}\,p\right) = 0$$

may be written as

$$\partial_t\,p + \partial_i\left(p\left(b_i - \frac{1}{2}\,\partial_j\left(\sigma_{ik}\sigma_{jk}\right)\right)\right) - \frac{1}{2}\,\partial_i\left(\sigma_{ik}\sigma_{jk}\partial_j\ p\right) = 0$$

that is

$$\partial_t\,p + \partial_i\left(pb^\sigma_i\right) - \frac{1}{2}\,\partial_i\left(\sigma_{ik}\sigma_{jk}\partial_j\ p\right) = 0. \tag{7.2}$$

This is equation (5.8) with $\mathbf{b}^\sigma$ instead of $\mathbf{b}$.

Likewise, we remark that the second order operator of the backward Kolmogorov equation (5.3)

$$\partial_t\,f\ - b_i\partial_i\,f - \frac{1}{2}\,\sigma_{ik}\sigma_{jk}\partial^2_{ij}\,f = 0.$$

may be written as

$$-\frac{1}{2}\,\sigma_{ik}\sigma_{jk}\partial^2_{ij}\,f = -\frac{1}{2}\,\partial_i\left(\sigma_{ik}\sigma_{jk}\partial_j\ f\right) + \frac{1}{2}\,\partial_i\left(\sigma_{ik}\sigma_{jk}\right)\partial_j\,f. \tag{7.3}$$

Thus (5.3) also reads as

$$\partial_t\,f - \left(b_i - \frac{1}{2}\,\partial_j\left(\sigma_{ik}\sigma_{jk}\right)\right)\partial_i\,f - \frac{1}{2}\,\partial_i\left(\sigma_{ik}\sigma_{jk}\partial_j\ f\right) = 0,$$

that is

$$\partial_t\,f - b^\sigma_i\,\partial_i\,f - \frac{1}{2}\,\partial_i\left(\sigma_{ik}\sigma_{jk}\partial_j\ f\right) = 0. \tag{7.4}$$

This latter equation is the adjoint of equation (5.8), where the drift vector $\mathbf{b}$ has been replaced by $\mathbf{b}^\sigma$. This can also be proved directly from (7.2) using that (5.3) and (5.2) are adjoint from one another.

It follows from the above observations that our results readily apply to (5.2), and, up to the passage to adjoint, to (5.3). We thus have.

Proposition 5 [Fokker-Planck equation] *Assume* $\mathbf{b}$ *and* σ *satisfy:*

$$\mathbf{b}^\sigma \in \Big(L^1([0,T], W^{1,1}_{loc}(\mathbb{R}^N))\Big)^N, \quad \operatorname{div}\mathbf{b}^\sigma \in L^1([0,T], L^\infty(\mathbb{R}^N)),$$
$$\frac{\mathbf{b}^\sigma}{1+|x|} \in \big(L^1([0,T], L^1 + L^\infty(\mathbb{R}^N))\big)^N \tag{7.5}$$

and (6.12). Then, for each initial condition in $L^1 \cap L^\infty$ *(resp.* $L^2 \cap L^\infty$*), the Fokker-Planck equation (5.2) has a unique solution in the space*

$$\Big\{ p \in L^\infty([0,T], L^1 \cap L^\infty) \quad (\text{resp.} \quad L^\infty([0,T], L^1 \cap L^\infty)),$$
$$\sigma^t \nabla p \in L^2([0,T], L^2) \Big\}.$$

Corollary 1 [Backward Kolmogorov equation] *The same conclusion holds* mutatis mutandis *for equation (5.3).*

Remark 18 *It is unclear to us how to interpret the "natural" assumption appearing in Proposition 5 on the* modified *drift vector* $\mathbf{b}^\sigma$ *in terms of the theory of stochastic differential equations. Why is it that this assumption does not* spontaneously *hold on the original drift vector* $\mathbf{b}$ *nor on the Stratanovich drift vector* $\mathbf{b}^{\text{Strat}}$ *(even if of course it can be translated in terms of the latter). This question could be addressed in the light of the relation between ordinary differential equations and stochastic differential equations (see [8, 9, 26, 27]).*

Remark 19 *To follow up on the previous remark, let us emphasize that we are not able to prove the results of Proposition 5 when* $\mathbf{b}$*, instead of* $\mathbf{b}^\sigma$*, satisfies (7.5). See however the two arguments in Section 8.1 that go in this direction.*

Of course, all the extensions of our results are still valid *modulo* the transformation of the drift vector by addition of $\sigma\theta$ as in Proposition 3.

An interesting extension is the following. We know from Proposition 3 that whenever we are able to treat $(\mathbf{b}, \sigma)$ we are also able to treat $(\mathbf{b} + \sigma\,\theta, \sigma)$, with $\theta \in L^2([0,T], L^2 + L^\infty(\mathbb{R}^N))^K$ (from (6.13)). So we may apply this to (5.2) written in the form (7.2). We therefore know Proposition 5 extends to any $\mathbf{b}^\sigma + \sigma\,\theta$ for convenient θ. Substracting (5.7) to (5.4), we see that

$$(\mathbf{b}^\sigma - \mathbf{b}^{\text{Strat}})_i = -\frac{1}{2}\,\sigma_{ik}\,\partial_j\,\sigma_{jk} = \sigma\,\theta,$$

with $\theta = -\frac{1}{2}\,\partial_j\,\sigma_{jk} \in L^2([0,T], L^2(\mathbb{R}^N))^K$ as soon as $\sigma \in L^2([0,T], W^{1,2} + W^{1,\infty}(\mathbb{R}^N))^{N\times K}$ Therefore, we may also conclude on the equation (5.2) with an assumptions on the drift vector $\mathbf{b}^{\text{Strat}}$ in place of those on $\mathbf{b}^\sigma$.

Of course, the same holds for (7.4)

We may also extend Proposition 4 (and the Remarks 16 and 17) to the setting of equations (5.2) and (5.3), simply replacing again $\mathbf{b}$ by $\mathbf{b}^\sigma$.

8 The Lipschitz case, and related issues

Assumption (7.5) states that $\mathbf{b}^\sigma = \mathbf{b} - \frac{1}{2}\,\mathrm{div}\,(\sigma\sigma^t)$ needs to be $W^{1,1}$, with its divergence controlled. Thus it implicitly contains a condition on second derivatives of σ (even if such derivatives are not *all* derivatives, and in addition are aggregated with derivatives of the drift vector $\mathbf{b}$). The question arises so as to know whether the existence and uniqueness result also holds for (5.2) and (5.3) in the absence of any such condition involving any second derivatives of σ. Of course, this is related to the treatment of the original drift vector $\mathbf{b}$ as noticed in Remarks 18 and 19.

In particular, it is of interest to test our techniques of proof on the classical case of Lipschitz regular coefficients b_i and σ_{ik}. In principle, this case is "the best case scenario", as recalled in Section 3.2.2. Our arguments above do not cover this case, precisely because of the need for some control on some second order derivatives of σ . Section 8.2 below show the type of results we are nevertheless able to establish in that situation, slightly modifying other assumptions (on the initial condition in particular).

Before we get to this Lipschitz case, we make in Section 8.1 several remarks on how to avoid assumptions on second order derivatives. The first remark, contained in the next section, turns out to be useful for the Lipschitz case, although it is important to note that this estimate is *not* restricted to the Lipschitz setting.

8.1 First order derivatives of σ

8.1.1 A specific version of the a priori estimate

First, we concentrate on the term that transforms (5.8) into (5.2), that is the last term of (7.1):

$$-\frac{1}{2}\,\partial_i\,(\partial_j\,(\sigma_{ik}\sigma_{jk})\,p).$$

A simple calculation shows that the corresponding term in the *a priori* estimate may be decomposed in the following manner:

$$\begin{aligned}\frac{1}{2}\int \partial_j\,(\sigma_{ik}\sigma_{jk})\,\partial_i\,\frac{f^2}{2} \quad &= \quad \int (\partial_i\,\sigma_{ik})\,(\sigma_{jk}\partial_j\,f)\,f \\ &\quad +\frac{1}{4}\int \big((\partial_i\,\sigma_{ik})^2 - (\partial_j\,\sigma_{ik})(\partial_i\,\sigma_{jk})\big)\,f^2. \end{aligned} \tag{8.1}$$

Indeed, we have, integrating by parts,

$$\frac{1}{2}\int (\partial_j\,\sigma_{ik})\sigma_{jk}\,\partial_i\,\frac{f^2}{2} = -\frac{1}{2}\int (\partial_i\,\sigma_{jk})\,(\partial_j\,\sigma_{ik})\,\frac{f^2}{2} - \frac{1}{2}\int \sigma_{jk}\,\partial^2_{ij}\,\sigma_{ik}\,\frac{f^2}{2}, \tag{8.2}$$

and likewise

$$\frac{1}{2}\int \sigma_{ik}(\partial_j\,\sigma_{jk})\,\partial_i\,\frac{f^2}{2} = -\frac{1}{2}\int (\partial_i\,\sigma_{ik})\,(\partial_j\,\sigma_{jk})\,\frac{f^2}{2} - \frac{1}{2}\int \sigma_{ik}\,\partial^2_{ij}\,\sigma_{jk}\,\frac{f^2}{2}, \tag{8.3}$$

Substracting (8.3) from (8.2) yields

$$\begin{aligned}\frac{1}{2}\int(\partial_j\,\sigma_{ik})\sigma_{jk}\,\partial_i\,\frac{f^2}{2} \quad &= \quad \frac{1}{2}\int\sigma_{ik}(\partial_j\,\sigma_{jk})\,\partial_i\,\frac{f^2}{2}\\ &\quad -\frac{1}{2}\int\left((\partial_i\,\sigma_{jk})\,(\partial_j\,\sigma_{ik})-(\partial_i\,\sigma_{ik})^2\right)\,\frac{f^2}{2}.\end{aligned} \tag{8.4}$$

Equation (8.1) is readily obtained adding $\displaystyle\int(\partial_j\,\sigma_{ik})\sigma_{jk}\,\partial_i\,\frac{f^2}{2}$ to both sides.

It follows that, from (5.2)

$$\partial_t\,p+\partial_i\,(pb_i)-\frac{1}{2}\,\partial^2_{ij}\,(\sigma_{ik}\sigma_{jk}\,p)=0,$$

the following L^2 energy estimate is obtained

$$\begin{aligned}&\frac{d}{dt}\int\frac{f^2}{2}+\int(\operatorname{div}\mathbf{b}\,)\frac{f^2}{2}+\int|\sigma^t\,\nabla f|^2\\ &\qquad -\int(\partial_i\,\sigma_{ik})\,(\sigma_{jk}\partial_j\,f)\,f-\frac{1}{4}\int\left((\partial_i\,\sigma_{ik})^2-(\partial_j\,\sigma_{ik})(\partial_i\,\sigma_{jk})\right)f^2=0,\end{aligned} \tag{8.5}$$

and this form is now preferred to (5.9) with $\mathbf{b}\ =\ \mathbf{b}^\sigma$. Each of the two new terms is now treated as follows. The first term may be bounded as follows

$$\left|\int(\partial_i\,\sigma_{ik})\,(\sigma_{jk}\partial_j\,f)\,f\right|\leq\|\partial_i\,\sigma_{ik}\|_{L^2(\mathbf{R}^N)}\left\|\sigma^t\nabla f\right\|_{L^2(\mathbf{R}^N)}\|f\|_{L^\infty(\mathbf{R}^N)} \tag{8.6}$$

while the second term is bounded by

$$\left|\int\left((\partial_i\,\sigma_{ik})^2-(\partial_j\,\sigma_{ik})(\partial_i\,\sigma_{jk})\right)f^2\right|\leq 2\,\|\partial_i\,\sigma_{ik}\|^2_{L^2(\mathbf{R}^N)}\|f\|^2_{L^\infty(\mathbf{R}^N)} \tag{8.7}$$

Consequently, assuming *e.g.*

$$\sigma\in L^2([0,T],W^{1,2}(\mathbb{R}^N))^{N\times K},$$

we obtain, say,

$$\frac{d}{dt}\int\frac{f^2}{2}+\int(\operatorname{div}\mathbf{b}\,)\frac{f^2}{2}+\frac{1}{2}\int|\sigma^t\,\nabla f|^2\leq C \tag{8.8}$$

for some constant C depending on the data. The conclusions of the *a priori* integrability of the solution, namely:

$$p\in L^\infty([0,T],L^2\cap L^\infty),\quad \sigma^t\nabla p\in L^2([0,T],L^2)$$

are not modified (for any *finite* final time T of course). A similar argument holds when

$$\sigma \in L^2([0,T], W^{1,2} + W^{1,\infty}(\mathbb{R}^N))^{N\times K}.$$

No condition on the second derivatives is required at this stage.

Unfortunately, we are not able to perform the regularization in the same manner, under the only first order assumption above. However, the above argument and the estimate will play a role in the case of Lipschitz regular coefficients we will consider in the next section.

Remark 20 *Note that all the above work in (8.1) to (8.4) aims to accomodate the fact that σ is* not *symmetric. Otherwise, things are much simpler.*

8.1.2 On the one-dimensional setting

A second remark that is in order concerns the one-dimensional setting. It is again related to the issue of symmetry, as Remark 20 above. When the ambient dimension N is $N = 1$, then the last term of (7.1) simplifies into $-\dfrac{1}{2}\,\partial_i\left(\partial_j\left(\sigma_{ik}\sigma_{jk}\right)p\right) = -\left(\sigma_{1k}\sigma'_{1k}\,p\right)'$, where the prime denotes differentiation with respect to the only space variable x_1. In other terms, the difference in drift vectors $\mathbf{b}^\sigma - \mathbf{b}$ is simply $\sigma_{1k}\sigma'_{1k}$. It is of the general form $\sigma\,\theta$ with $\theta_k = \sigma'_{1k}$ square integrable in space as soon as $\sigma \in L^2([0,T], W^{1,2}(\mathbb{R}))^K$. The same argument applies to the Stratanovich drift vector $\mathbf{b}^{\mathrm{Strat}}$. Consequently, we deduce from Propositions 2, 3, and 5:

Corollary 2 [One-dimensional setting] *In one dimension, assume that any one of the three drift vector $\mathbf{b}$, $\mathbf{b}^\sigma$ or $\mathbf{b}^{\mathrm{Strat}}$ satisfies*

$$\begin{aligned} \mathbf{b} \in \left(L^1([0,T], W^{1,1}_{loc}(\mathbb{R}))\right), \quad \operatorname{div}\mathbf{b} \in L^1([0,T], L^\infty(\mathbb{R})), \\ \frac{\mathbf{b}}{1+|x|} \in \left(L^1([0,T], L^1 + L^\infty(\mathbb{R}))\right) \end{aligned} \tag{8.9}$$

and that σ satisfies

$$\sigma \in \left(L^2([0,T], W^{1,2}_{loc}(\mathbb{R}))\right)^K, \quad \frac{\sigma}{1+|x|} \in \left(L^2([0,T], L^2 + L^\infty(\mathbb{R}))\right)^K \tag{8.10}$$

Then, for each initial condition in $L^1 \cap L^\infty$ (resp. $L^2 \cap L^\infty$), the Fokker Planck equation of divergence form (5.8), the Fokker-Planck equation (5.2), and the backward Kolmogorov equation (5.3) all have a unique solution in the space

$$\Big\{ p \in L^\infty([0,T], L^1 \cap L^\infty) \quad (\text{resp.} \quad L^\infty([0,T], L^1 \cap L^\infty)), \\ \sigma^t \nabla p \in L^2([0,T], L^2) \Big\}.$$

8.1.3 On the one-dimensional setting again

Wee would now like to comment on the $W^{1,2}$ Sobolev regularity we impose on the dispersion matrix σ. We may provide one argument in order to show some optimality of this assumption. Again, this argument concerns the one-dimensional setting, and we are not able to extend it to higher dimensions.

We anticipate on the arguments of [18] and momentarily admit that the results we have proved here on Fokker-Planck type equations have immediate consequences on the well-posedness of the stochastic differential equations. More precisely, we admit that the assumption $\sigma \in W^{1,2}$ allows for proving pathwise uniqueness for the stochastic differential equation (using the Fokker-Planck equation in doubled dimension, as introduced in Section 3.3)

Then we remark that the limit case for the Sobolev imbedding in one dimension precisely reads $W^{1,2} \hookrightarrow C^{0,1/2}$. As a matter of fact, the $C^{0,1/2}$ Schauder regularity is the "natural" regularity that allows to show pathwise uniqueness (and is actually the limit case for this, but we will not show this sharpness here). The result is well-known (see e.g. [14, 15, 25]). One possible argument is the following.

Let us consider two (one-dimensional) solutions X_t and Y_t of the stochastic differential equation (3.3) with the same (K-dimensional) brownian motion $\mathbf{W}_t$:

$$dX_t = \sigma(X_t)\, d\mathbf{W}_t, \quad dY_t = \sigma(Y_t)\, d\mathbf{W}_t,$$

and the same initial condition. Note that we have assume there is no drift: $\mathbf{b} = 0$. We next introduce for $t \geq 0$ the cut-off function

$$\psi_\varepsilon(t) = \begin{cases} t \log\left(\frac{t}{\varepsilon}\right) - (t-\varepsilon) & \text{when} \quad t \geq \varepsilon \\ 0 & \text{otherwise} \end{cases} \tag{8.11}$$

It is simple to see that this function is convex and twice differentiable, with

$$\psi_\varepsilon''(t) = \begin{cases} \frac{1}{t} & \text{when} \quad t \geq \varepsilon \\ 0 & \text{otherwise.} \end{cases} \tag{8.12}$$

By the Ito rule, we readily have

$$\begin{aligned} &\psi_\varepsilon(|X_t - Y_t|) - \psi_\varepsilon(0) \\ &\qquad = \int_0^t \psi_\varepsilon'(|X_s - Y_s|)\ (\sigma(X_s) - \sigma(Y_s))\ d\mathbf{W}_t \\ &\qquad +\frac{1}{2}\int_0^t \psi_\varepsilon''(|X_s - Y_s|)\ (\sigma(X_s) - \sigma(Y_s))\ \big(\sigma^t(X_s) - \sigma^s(Y_s)\big)\ ds \end{aligned} \tag{8.13}$$

Thus,

$$\begin{aligned} &\psi_\varepsilon(|X_t - Y_t|) \\ &\qquad -\frac{1}{2}\int_0^t \psi_\varepsilon''(|X_s - Y_s|)\ (\sigma(X_s) - \sigma(Y_s))\ \big(\sigma^t(X_s) - \sigma^t(Y_s)\big)\ ds \\ &\qquad\quad = M_t \end{aligned} \tag{8.14}$$

is a martingale. We now fix $\delta > 0$. Denoting by τ the stopping time

$$\tau = \inf\{t \geq 0, \quad |X_t - Y_t| \geq \delta\},$$

we deduce from (8.14) and the stopping time theorem that

$$\mathbb{E}\Bigg(\psi_\varepsilon(|X_{t\wedge\tau} - Y_{t\wedge\tau}|) \\ -\frac{1}{2}\int_0^{t\wedge\tau} \psi_\varepsilon''(|X_s - Y_s|)\,(\sigma(X_s) - \sigma(Y_s))\,\big(\sigma^t(X_s) - \sigma^t(Y_s)\big)\,ds\Bigg) \\ = \mathbb{E}(M_{t\wedge\tau}) = \mathbb{E}(M_0) = 0. \tag{8.15}$$

Since ψ_ε'' is given by (8.12), we have

$$0 \leq \psi_\varepsilon''(z) \leq \frac{1}{z}$$

for all $t \geq 0$, and thus in particular for $z = |X_s - Y_s|$. On the other hand, σ is assumed $C^{0,1/2}$ thus

$$\Big|(\sigma(X_s) - \sigma(Y_s))\,\big(\sigma^t(X_s) - \sigma^t(Y_s)\big)\Big| \leq C\,|X_s - Y_s|,$$

for some irrelevant constant C. Therefore (8.15) reads

$$\mathbb{E}\left(\psi_\varepsilon(|X_{t\wedge\tau} - Y_{t\wedge\tau}|)\right) \\ = \mathbb{E}\left(\frac{1}{2}\int_0^{t\wedge\tau} \psi_\varepsilon''(|X_s - Y_s|)\,(\sigma(X_s) - \sigma(Y_s))\,\big(\sigma^t(X_s) - \sigma^t(Y_s)\big)\,ds\right) \\ \leq C\,t\wedge\tau \leq C\,t \tag{8.16}$$

This implies that

$$\mathbb{P}(\tau \leq t)\,\psi_\varepsilon(\delta) = \mathbb{P}(\tau \leq t)\,\psi_\varepsilon(|X_\tau - Y_\tau|) \leq \mathbb{E}\left(\psi_\varepsilon(|X_{t\wedge\tau} - Y_{t\wedge\tau}|)\right) \leq C\,t.$$

Letting ε go to zero, we have $\psi_\varepsilon(\delta) \longrightarrow +\infty$ and thus, keeping t and δ fixed, we obtain $\mathbb{P}(\tau \leq t) = 0$ for all times t. This shows $\mathbb{P}(\tau < +\infty) = 0$. Since this holds for all δ, we obtain $X_t = Y_t$ and pathwise uniqueness.

8.2 Lipschitz regular coefficients

We are now in position to consider Lipschitz regular coefficients $\mathbf{b}$ and σ. In fact in this specific case, the technique of proof, and the results themselves, sensitively depend on the regularity of the initial condition. The cases of a H^1, a Lipschitz regular, and a L^2 initial condition will be examined in the three following subsections, respectively.

At this stage of the exposition, the reader is assumed to be familiar with the basic techniques and ingredients of our work. We will therefore sketch the proof of the results, only making explicit the significant differences with the previously considered settings.

8.2.1 H^1 initial conditions

For H^1 initial conditions, we are able to prove the.

Proposition 6 [Lipschitz regular coefficients, H^1 initial condition] *Assume* $\mathbf{b}$ *and* σ *are time-independent and satisfy:*

$$\mathbf{b} \in \left(Lip_{loc}(\mathbb{R}^N)\right)^N, \quad \operatorname{div}\mathbf{b} \in L^\infty(\mathbb{R}^N), \quad \frac{\mathbf{b}}{1+|x|} \in \left(L^1 + L^\infty(\mathbb{R}^N)\right)^N$$

$$\sigma \in \left(Lip(\mathbb{R}^N)\right)^{N\times K}, \quad \frac{\sigma}{1+|x|} \in \left(L^2 + L^\infty(\mathbb{R}^N)\right)^{N\times K},$$

along with the additional assumption

$$\frac{\partial_i b_j + \partial_j b_i}{2} \leq C\,\mathrm{Id}, \tag{8.17}$$

in the sense of symmetric matrices, for some constant C *independent of* x*. Then, for each initial condition in* H^1*, equation (2.12), that is*

$$\partial_t f - b_i \partial_i f - \frac{1}{2}\sigma_{ik}\sigma_{jk}\partial^2_{ij} f = 0$$

has a unique solution in the space $L^\infty([0,T], H^1)$*.*

Remark 21 *Of course the usual extensions are possible:* $\operatorname{div}\mathbf{b}$ *bounded from below,* L^1 *time-dependence, etc. The case of a definite positive matrix* $\sigma^t\sigma$ *is of course simpler. Also, the adjoint form of the equation may be considered. We omit all this here.*

To prove the uniqueness statement contained in Proposition 6 (the existence statement is, as always, the easy part), the key step is the *a priori* estimate, which basically consists in proving that the H^1 regularity (in space) of the solution at initial time is propagated forward in time, so that the natural solution space is $L^\infty([0,T], H^1)$. Once this formal estimate is established, the regularization step is straightforward. It is indeed well known that when the solution f is H^1, a transport term $\mathbf{c}\cdot\nabla f$ may be adequately regularized with $\mathbf{b}$ only in L^p_{loc}, say. No $W^{1,1}$ Sobolev regularity is then needed on $\mathbf{c}$. This is the last statement of Lemma 1 recalled above. Applying this statement to $\mathbf{c} = \mathbf{b} - \frac{1}{2}\partial_j\left(\sigma_{ik}\sigma_{jk}\right)$, we obtain, for Lipschitz $\mathbf{b}$ and σ, that $\mathbf{c}$ is locally bounded and thus $[\rho_\varepsilon, \mathbf{c}\cdot\nabla](f)$ vanishes in L^2_{loc} with ε. The other terms of the equation are regularized in the usual way, as performed in Section 5.3. Uniqueness follows.

So let us outline the proof of the adequate formal H^1 *a priori* estimate.

To begin with, we differentiate equation (2.12) with respect to the m-th space variable x_m:

$$\partial_t\,\partial_m f - b_i\partial^2_{im} f - \left(\partial_m\, b_i\right)\partial_i f - \frac{1}{2}\partial_m\left(\sigma_{ik}\sigma_{jk}\right)\partial^2_{ij} f - \frac{1}{2}\sigma_{ik}\sigma_{jk}\partial_m\,\partial^2_{ij} f = 0,$$

next multiply by $\partial_m f$, sum over m and integrate over the space. This leads to:

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\int |\nabla f|^2 + \frac{1}{2}\int \operatorname{div}\mathbf{b}\ |\nabla f|^2 - \int (\partial_m b_i)\, \partial_i f\, \partial_m f \\ -\frac{1}{2}\int \sigma_{ik}\sigma_{jk}\partial_m\, \partial^2_{ij} f \partial_m f - \frac{1}{2}\int \partial_m (\sigma_{ik}\sigma_{jk})\, \partial^2_{ij} f\, \partial_m f = 0.\end{aligned} \tag{8.18}$$

The first two terms are standard. Note that for the second one, only a control *from below* on $\operatorname{div}\mathbf{b}$, namely (5.12), is needed, as pointed out in Remarks 3 and 21. The third one can be taken care of, assuming (8.17), which implies

$$-\int (\partial_m b_i)\, \partial_i f\, \partial_m f \geq -C\int |\nabla f|^2. \tag{8.19}$$

The fourth term is treated as follows:

$$\begin{aligned}-\frac{1}{2}\int \sigma_{ik}\sigma_{jk}\partial_m\, \partial^2_{ij} f \partial_m f &= \frac{1}{2}\int \sigma_{ik}\sigma_{jk}\partial^2_{im} f \partial^2_{jm} f \\ &+ \frac{1}{2}\int \partial_j (\sigma_{ik}\sigma_{jk})\partial^2_{im} f \partial_m f.\end{aligned} \tag{8.20}$$

We henceforth denote by

$$\int \left|\sigma^t\nabla(\nabla f)\right|^2 = \int \sum_k \sum_m \left(\sigma_{ik}\, \partial^2_{im} f\right)^2. \tag{8.21}$$

We then remark, using for each m the identity (8.1) with $\partial_m f$ instead of f, and summing over m,

$$\begin{aligned}\frac{1}{2}\int \partial_j (\sigma_{ik}\sigma_{jk})\partial^2_{im} f \partial_m f &= \int (\partial_i\, \sigma_{ik})\, (\sigma_{jk}\partial^2_{jm} f)\, \partial_m f \\ &+ \frac{1}{4}\int \left((\partial_i\, \sigma_{ik})^2 - (\partial_j\, \sigma_{ik})(\partial_i\, \sigma_{jk})\right) |\nabla f|^2.\end{aligned}$$

Inserting this in (8.20), we obtain the new expression for the fourth term of (8.18):

$$\begin{aligned}&-\frac{1}{2}\int \sigma_{ik}\sigma_{jk}\partial_m\, \partial^2_{ij} f \partial_m f \\ &\quad = \frac{1}{2}\int \left|\sigma^t\nabla(\nabla f)\right|^2 + \int (\partial_i\, \sigma_{ik})\, (\sigma_{jk}\partial^2_{jm} f)\, \partial_m f \\ &\quad\quad + \frac{1}{4}\int \left((\partial_i\, \sigma_{ik})^2 - (\partial_j\, \sigma_{ik})(\partial_i\, \sigma_{jk})\right) |\nabla f|^2.\end{aligned} \tag{8.22}$$

and thus, using the Young inequality, the estimation

$$-\frac{1}{2}\int \sigma_{ik}\sigma_{jk}\partial_m\, \partial^2_{ij} f \partial_m f \geq \left(\frac{1}{2} - \eta\right)\int \left|\sigma^t\nabla(\nabla f)\right|^2 - C_\eta \left\|\nabla\sigma\right\|^2_{L^\infty} \int |\nabla f|^2, \tag{8.23}$$

for η small and some irrelevant constant C_η. It remains now to treat the fifth term of (8.18), indeed by the same technique:

$$\begin{aligned}
-\frac{1}{2}\int \partial_m\,(\sigma_{ik}\sigma_{jk})\,\partial^2_{ij}\,f\,\partial_m\,f \quad &= \quad -\frac{1}{2}\int (\partial_m\,\sigma_{ik})\,(\sigma_{jk}\,\partial^2_{ij}\,f)\,\partial_m\,f \\
&\qquad -\frac{1}{2}\int (\partial_m\,\sigma_{jk})\,(\sigma_{ik}\,\partial^2_{ij}\,f)\,\partial_m\,f \\
&\geq \quad -\eta\int \left|\sigma^t\nabla(\nabla f)\right|^2 - C_\eta\,\|\nabla\sigma\|^2_{L^\infty}\int |\nabla f|^2 .
\end{aligned}$$

Collecting (8.19) to (8.24) and inserting all of them in (8.18), we obtain the energy estimate:

$$\frac{1}{2}\frac{d}{dt}\int |\nabla f|^2 + \frac{1}{4}\int \left|\sigma^t\nabla(\nabla f)\right|^2 \leq \mathcal{C}\int |\nabla f|^2 , \tag{8.24}$$

for some “large” constant $\mathcal{C}$, depending on the various previous constants, C in (8.17), $\|\mathrm{div}\,\mathbf{b}\,\|_{L^\infty}$, and $\|\nabla\sigma\|_{L^\infty}$.

Next, we consider the formal L^2 estimate (8.5) obtained above, namely

$$\begin{aligned}
&\frac{d}{dt}\int \frac{f^2}{2} + \int (\mathrm{div}\,\mathbf{b}\,)\frac{f^2}{2} + \int |\sigma^t\,\nabla f|^2 \\
&\quad - \int (\partial_i\,\sigma_{ik})\,(\sigma_{jk}\partial_j\,f)\,f - \frac{1}{4}\int \big((\partial_i\,\sigma_{ik})^2 - (\partial_j\,\sigma_{ik})(\partial_i\,\sigma_{jk})\big)\,f^2 = 0,
\end{aligned}$$

which gives

$$\frac{1}{2}\frac{d}{dt}\int f^2 + \frac{1}{2}\int |\sigma^t\,\nabla f|^2 \leq \mathcal{C}\int f^2, \tag{8.25}$$

again for some constant $\mathcal{C}$ depending on $\|\mathrm{div}\,\mathbf{b}\,\|_{L^\infty}$, and $\|\nabla\sigma\|_{L^\infty}$.

The combination of (8.24) and (8.25) shows that, formally,

$$f \in L^\infty([0,T], H^1),$$

and the proof can be completed.

Remark 22 *In addition to the $L^\infty([0,T],H^1)$ regularity, we also get from the above proof*

$$\int_0^T\int |\sigma^t\,\nabla f|^2 + \int_0^T\int \left|\sigma^t\nabla(\nabla f)\right|^2 < \infty.$$

Remark 23 *In fact, we have made use in the above proof of the* global *Lipschitz regularity of σ. See indeed equations (8.24) and (8.25). The proof might be adapted to cover the case when σ is only* locally *Lipschitz, at the price of putting some extra growth assumptions on $\nabla\sigma$ at infinity. On the other hand, the* local *Lipschitz regularity of $\mathbf{b}$ is sufficient, because we have the global controls on $\|\mathrm{div}\,\mathbf{b}\,\|_{L^\infty}$ and (8.17).*

8.2.2 Lipschitz initial conditions

It is well known that if the coefficients $\mathbf{b}$ and σ are Lipschitz regular, and if the solution of the Fokker-Planck equation (or that of the bacward Kolmogorov equation) is Lipschitz at initial time, then it is Lipschitz for all time. One way of proving this, say for the equation (2.12), is to consider the representation formula for the solution, in terms of the solution of the stochastic differential equation. The Lipschitz regularity follows.

We are now going to prove this propagation of the Lipschitz regularity, arguing only on the partial differential equation. This will allow us to circumvent the question of relating the partial differential equation to the stochastic differential equation, and also to prove an existence and uniqueness result in the appropriate class of functions. In other words, we know from probability theory that in this setting ($\mathbf{b}$, σ and the initial condition all Lipschitz), there exists at least one solution to the PDE, which is Lipschitz regular for all times. Does the solution constructed from our *a priori + regularization* procedure coincide with that one? The answer turns out to be positive, and this is the purpose of:

Proposition 7 [Lipschitz regular coefficients, H^1 initial condition]
Assume the conditions on $\mathbf{b}$ and σ contained in the statement of Proposition 6. Assume that the initial condition f_0 is Lipschitz. Then equation equation (2.12)

$$\partial_t f - b_i \partial_i f - \frac{1}{2}\sigma_{ik}\sigma_{jk}\partial^2_{ij} f = 0$$

has a unique solution Lipschitz regular in space, for all times. In addition, this unique solution satisfies the L^2 estimate

$$\frac{d}{dt}\int f^2 + \int |\sigma^t \nabla f|^2 \leq \mathcal{C} \int f^2. \tag{8.26}$$

The proof of the above proposition relies on arguments that are now well known to the reader. The only point we have to make is that the Lipschitz regularity is propagated forward in time, that is, $f(t,x)$ solution to (2.12) is Lipschitz-in-x if f_0 is.

The formal proof of the Lipschitz regularity is a consequence of the maximum principle for equation (2.12). We compute $\partial_t |\nabla f|^2$ using the expression of $\partial_t f$ provided by the equation:

$$\begin{aligned}
\frac{1}{2}\partial_t (\partial_m f)^2 &= \partial_m f\, \partial^2 tmf \\
&= \partial_m f\, b_i \partial^2_{im} f + \partial_m f\, \partial_m b_i \partial_i f + \frac{1}{2}\partial_m f\, \sigma_{ik}\sigma_{jk}\partial_m \partial^2_{ij} f \\
&\quad + \frac{1}{2}\partial_m f\, \partial_m (\sigma_{ik}\sigma_{jk})\partial^2_{ij} f.
\end{aligned} \tag{8.27}$$

We successively treat each term of the right-hand side. First,

$$\partial_m f\, b_i \partial^2_{im} f = \mathbf{b} \cdot \nabla \left(\frac{|\nabla f|^2}{2}\right) \tag{8.28}$$

Second, using (8.17),

$$|\partial_m f \, \partial_m b_i \partial_i f| \leq C |\nabla f|^2. \tag{8.29}$$

Third,

$$\begin{aligned} \partial_m f \, \sigma_{ik}\sigma_{jk}\partial_m\partial^2_{ij} f &= \sigma_{ik}\sigma_{jk}\,\partial^2_{ij}\left(\frac{|\nabla f|^2}{2}\right) - \sigma_{ik}\sigma_{jk}\partial^2_{im} f \, \partial^2_{jm} f \\ &= \sigma_{ik}\sigma_{jk}\,\partial^2_{ij}\left(\frac{|\nabla f|^2}{2}\right) - \left|\sigma^t\nabla(\nabla f)\right|^2 . \end{aligned} \tag{8.30}$$

using notation (8.21). And eventually

$$\begin{aligned} \left|\partial_m f \, \partial_m \left(\sigma_{ik}\sigma_{jk}\right)\partial^2_{ij} f\right| &= \left|\partial_m f \left(\partial_m \sigma_{ik}\right)\sigma_{jk}\partial^2_{ij} f + \partial_m f \left(\partial_m \sigma_{jk}\right)\sigma_{ik}\partial^2_{ij} f\right| \\ &\leq |\nabla f| \; \|\nabla\sigma\|_{L^\infty} \left|\sigma^t\nabla(\nabla f]\right| \end{aligned} \tag{8.31}$$

Collecting all the above, we obtain from (8.27):

$$\begin{aligned} &\partial_t\left(\frac{|\nabla f|^2}{2}\right) - \mathbf{b}\cdot\nabla\left(\frac{|\nabla f|^2}{2}\right) - \frac{1}{2}\sigma_{ik}\sigma_{jk}\,\partial^2_{ij}\left(\frac{|\nabla f|^2}{2}\right) + \left|\sigma^t\nabla(\nabla f)\right|^2 \\ &\quad \leq |\nabla f| \; \|\nabla\sigma\|_{L^\infty} \left|\sigma^t\nabla(\nabla f]\right| \end{aligned} \tag{8.32}$$

from where we deduce

$$\partial_t\left(\frac{|\nabla f|^2}{2}\right) - \mathbf{b}\cdot\nabla\left(\frac{|\nabla f|^2}{2}\right) - \frac{1}{2}\sigma_{ik}\sigma_{jk}\,\partial^2_{ij}\left(\frac{|\nabla f|^2}{2}\right) \leq \mathcal{C}\;|\nabla f|\,, \tag{8.33}$$

and using the maximum principle, the Lipschitz regularity of f for all times.

Now that we have established the (formal) Lipschitz regularity of the solution, we argue as follows. We regularize all the data so as to obtain a smooth solution. This smooth solution satisfies the L^2 estimate, and satisfies also the Lipschitz regularity. Both facts are now rigorous, and not only formal, and hold independently of the regularization parameter ε. As ε goes to zero, we obtain a solution to the equation, which still satisfies these two conditions. This shows the existence in the appropriate class. On the other hand, uniqueness is also easy: the difference of two such solutions is Lipschitz, and thus the regularization step readily carries through.

8.2.3 L^2 initial conditions

We conclude this section, and this article, with the case of a L^2 initial condition, and Lipschitz regular $\mathbf{b}$ and σ.

First, we may approximate the L^2 initial condition f_0 by a sequence f_0^n of Lipschitz functions, converging in L^2 to f_0. Each of these f_0^n may be taken as an initial condition for the equation, giving rise to some solution $f_n(t,x)$, agreeing at initial time with f_0^n. All functions f_n are Lipschitz and satisfy estimate (8.5). It is a straightforward consequence of the arguments and results of the preceeding section that the sequence of functions f_n is converging to some f, which solves the equation with initial condition f_0. Remark indeed that since we may regularize Lipschitz solutions, we may estimate the distance between two such solutions, and prove that the sequence of solutions is a Cauchy sequence in $L^\infty([0,T],L^2)$, and $\sigma^t\nabla f_n$ is also a Cauchy sequence in $L^2([0,T],L^2)$ (because, at initial time, f_0^n is a Cauchy sequence in L^2). Therefore it converges, and, by the same argument, the limit is shown to be a solution, for initial condition f_0, which satisfies estimate (8.5). This defines the unique prolongation of Lipschitz solutions to L^2 initial conditions.

On the other hand, if we are now given a solution to (2.12) with L^2 initial condition f_0, it is not clear that this solution is this unique prolongation, and thus is unique. All what we are able to prove in this direction is the following.

Considering (8.5), let us denote by

$$\begin{aligned} E(T,f) &= \int \frac{f^2}{2}(T) + \int_0^T \int (\operatorname{div}\mathbf{b})\frac{f^2}{2} + \int_0^T \int |\sigma^t\,\nabla f|^2 \\ &\quad - \int_0^T \int (\partial_i\,\sigma_{ik})\,(\sigma_{jk}\partial_j\,f)\,f \\ &\quad - \frac{1}{4}\int_0^T \int \big((\partial_i\,\sigma_{ik})^2 - (\partial_j\,\sigma_{ik})(\partial_i\,\sigma_{jk})\big)\,f^2 \end{aligned} \tag{8.34}$$

We next introduce the following formulation of equation (2.12): f is said to be a solution *in the energy sense* to (2.12), with initial condition f_0 if, for all $0 \le t \le T$ and all smooth functions $\varphi(s,x)$, we have

$$\begin{aligned} E(t,f-\varphi) &= -\int_0^t \int \left(\partial_t\,\varphi - b_i\partial_i\,\varphi - \frac{1}{2}\sigma_{ik}\sigma_{jk}\partial_{ij}^2\,\varphi\right)(f-\varphi) \\ &\quad + \frac{1}{2}\int |f_0 - \varphi(0,\cdot)|^2 \end{aligned} \tag{8.35}$$

It may easily be verified that the above formal condition (8.35) amounts to

$$\left\{ \begin{array}{l} \partial_t\,f - b_i\partial_i\,f - \frac{1}{2}\sigma_{ik}\sigma_{jk}\partial_{ij}^2\,f = 0, \\ f(0) = f_0, \end{array} \right.$$

This owes to the fact that, formally again, because of (8.5) and (8.34),

$$\begin{aligned} E(t,f-\varphi) &= \int_0^t \int \left(\partial_t\, - b_i\partial_i\, - \frac{1}{2}\sigma_{ik}\sigma_{jk}\partial_{ij}^2\right)(f-\varphi)\,\cdot\,(f-\varphi) \\ &\quad + \frac{1}{2}\int |f(0,\cdot) - \varphi(0,\cdot)|^2 \end{aligned} \tag{8.36}$$

The condition (8.35) therefore embodies equation (2.12) without explicitly manipulating the latter.

Then we may show that, given a L^2 initial condition f_0, we may construct a solution to (2.12) in the energy sense (8.35), and that this solution is unique in this class.

Indeed, approximating f_0 by a sequence of Lipschitz functions, we simply consider the above defined unique prolongation of Lipschitz solutions, starting from f_0. As, again, Lipschitz solutions may be approximated by regularization, they satisfy (8.35). This shows existence.

As regards uniqueness, we argue as follows. Being given a solution g in the sense of (8.35), we compare it with the solution f obtained by Lipschitz approximation. For this purpose, consider f_n a sequence of Lipschitz solution that converges to f, and regularize further these solutions f_n by convolution, so as to obtain $f_{n,\varepsilon} = f_n \star \rho_\varepsilon$. Clearly, by the arguments of this article,

$$\left(\partial_t - b_i\partial_i - \frac{1}{2}\sigma_{ik}\sigma_{jk}\partial^2_{ij}\right) f_{n,\varepsilon} \longrightarrow 0,$$

in L^2, for n fixed, as ε goes to zero. We next use this smooth function $f_{n,\varepsilon}$ as a test function in (8.35), and obtain, letting ε go to zero:

$$E(t, g - f^n) = \frac{1}{2}\int |g(0) - f^n(0)|^2 .$$

Since $g(0) = f_0$, $f^n(0) = f_0^n$, and f_0^n converges to f_0 as n goes to infinity, we obtain $E(t, g - f) = 0$, and thus $g = f$. This shows the claimed uniqueness.

Remark 24 *We do not know whether existence and uniqueness of the solution may be established, for* $\mathbf{b}$ *and* σ *lipschitz,* $f_0 \in L^2 \cap L^\infty$ *(say), even when the solution is* assumed *to satisfy the energy estimate (8.5).*

References

[1] L. Ambrosio, Transport equation and Cauchy problem for BV vector fields, , *Inventiones Mathematicae*, 158 (2004), 227-260.

[2] L. Ambrosio, M. Lecumberry & S. Maniglia, Lipschitz regularity and approximate differentiability of the DiPerna-Lions flow, *Rendiconti del Seminario Fisico Matematico di Padova*, 114 (2005), 29-50.

[3] L. Ambrosio, G. Crippa, Existence, uniqueness, stability and differentiability properties of the flow associated to weakly differentiable vector fields, *preprint*, 2006.

[4] G. Barles, P.-L. Lions, Remarques sur les problèmes de réflexion oblique, [Remarks on oblique reflection problems], *C.R. Acad. Sc.*, t. 320, Série 1, 1995, pp 69-74.

[5] E. Carlen, M. Loss, Optimal smoothing and decay estimates for viscously damped conservation laws, with applications to the 2-D Navier-Stokes equation. A celebration of John F. Nash, Jr. , *Duke Math. J.* 81 (1995), no. 1, 135–157.

[6] R.J. Di Perna and P.L. Lions, Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.*, 98(3):511-547, 1989.

[7] R.J. Di Perna and P.L. Lions, On the Fokker-Planck-Boltzmann equation, *Commu. Math. Phys.*, 120, 1988, pp 1-23.

[8] H. Doss, Liens entre équations différentielles stochastiques et ordinaires, *Ann. IHP* Sec. B, vol 13, 1977, 2, pp 99-125.

[9] H. Doss, Liens entre équations différentielles stochastiques et ordinaires, *C.R. Acad. Sc. Paris*, Sér. A-B, vol 283, 1976, 13, A939-942.

[10] A. Figalli, Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients, *preprint*, October 2006.

[11] E. Fournié, J.-M. Lasry, J. Lebuchoux and P.-L. Lions, Applications of Malliavin calculus to Monte-Carlo methods in finance., *Finance Stoch.*, vol. 3, 4, 1999, pp 391-412.

[12] E. Fournié, J.-M. Lasry, J. Lebuchoux and P.-L. Lions, Applications of Malliavin calculus to Monte-Carlo methods in finance. II., *Finance Stoch.*, vol. 5, 2, 2001, pp 201-236.

[13] M. Hauray, C. Le Bris, P.L. Lions, Deux remarques sur les flots généralisés d'équations différentielles ordinaires, [Two remarks on generalized flows for ordinary differential equations], *C.R. Acad. Sc., Paris, Sér. I., Math.*, submitted.

[14] N. Ikeda, S. Watanabe, **Stochastic differential equations and diffusions processes**, 2nd Edition, North-Holland, 1989.

[15] I. Karatzas, S. Shreve, **Brownian motion and stochastic calculus**, Graduate texts in Mathematics, vol. 113, Springer, 1988.

[16] N. Lerner, Transport equations with partially BV velocities, *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) 3, no. 4, 681–703, 2004.

[17] C. Le Bris, P.L. Lions, Renormalized solutions of some transport equations with partially $W^{1,1}$ velocities and applications, *Annali di Matematica pura ed applicata*, volume 183, 97-130, 2004.

[18] C. Le Bris, P.L. Lions, Generalized flows for stochastic differential equations with irregular coefficients, in preparation.

[19] P.L. Lions, Sur les équations différentielles ordinaires et les équations de transport, [On ordinary differential equations and transport equations], *C.R. Acad. Sc., Paris, Sér. I., Math.*, 326(7):833-838, 1998.

[20] H. Osada, Diffusion processes with generators of generalized divergence form, *J. Math. Kyoto Univ.*, vol. 27, 1987, 4, pp 597-619.

[21] H-C. Ottinger, **Stochastic processes in polymeric fluids**, Springer, 1996.

[22] R. G. Owens and T. N. Phillips, **Computational Rheology**, Imperial College Press, 2002.

[23] D. Revuz, M. Yor, **Continuous martingales and brownian motion**, Grundlehren der mathematischen Wissenschaften. **293**, Springer, 3rd edition, 1999.

[24] L.C.G. Rogers, D. Williams, **Diffusions, Markov processes and martingales**, Cambridge Mathematical Library, 2volumes, 2nd edition, 2000.

[25] D. Stroock and S.R.S. Varadhan, **Multidimensional diffusion processes**, Grundlehren der mathematischen Wissenschaften. **233**, Springer, 1979.

[26] H.J. Sussmann, On the gap between deterministic and stochastic differential equations, *Ann. Probability*, vol. 6, 1978, 1, pp 19-41.

[27] H.J. Sussmann, An interpretation of stochastic differential equations as ordinary differential equations which depend on the sample point, *Bull. A. M. S.*, vol 83, 1977, 2, pp 296-298.