Degenerate McKean-Vlasov equations with drift in anisotropic negative Besov spaces

Elena Issoglio, Stefano Pagliarani, Francesco Russo, Davide Trevisani

Degenerate McKean-Vlasov equations with drift in anisotropic negative Besov spaces

Elena Issoglio * Stefano Pagliarani † Francesco Russo ‡ Davide Trevisani §

This version: 16th January 2024

Abstract

The paper is concerned with a McKean-Vlasov type SDE with drift in anisotropic Besov spaces with negative regularity and with degenerate diffusion matrix under the weak Hörmander condition. The main result is of existence and uniqueness of a solution in law for the McKean-Vlasov equation, which is formulated as a suitable martingale problem. All analytical tools needed are derived in the paper, such as the well-posedness of the Fokker-Planck and Kolmogorov PDEs with distributional drift, as well as continuity dependence on the coefficients. The solutions to these PDEs naturally live in anisotropic Besov spaces, for which we developed suitable analytical inequalities, such as Schauder estimates.

Keywords: McKean-Vlasov; kinetic PDEs; martingale problem; singular drift; anisotropic Besov-Hölder spaces; Schauder estimates.
MSC 2020: 60H10; 60H30; 60H50; 35C99; 35D99; 35K10.

Contents

*Dipartimento di Matematica G. Peano, Università di Torino, Torino, Italy. **E-mail**: elena.issoglio@unito.it
†Dipartimento di Matematica, Università di Bologna, Bologna, Italy. **E-mail**: stefano.pagliarani9@unibo.it.
‡Unité de Mathématiques Appliquées, ENSTA Paris, Institut Polytechnique de Paris, Palaiseau, France. **E-mail**: francesco.russo@ensta-paris.fr
§Facultad de Informática, Universitade da Coruña, A Coruña, Spain. **E-mail**: davide.trevisani@udc.es

1 Introduction and preliminaries

1.1 Statement of the problem

In this work we treat the weak well-posedness for a class of whose drift contains an anisotropic Besov distribution of negative order $-\beta$, and $\beta \in (0, 1/2)$. Let $0 < d < N$ be two given natural numbers, and let us denote by $Z_t := (V_t, X_t)$ a stochastic process valued on $\mathbb{R}^N = \mathbb{R}^d \times \mathbb{R}^{N-d}$.

The class of equations that we consider in this paper is given as

$$\begin{cases} dV_t = \Big(F\big(u_t(Z_t)\big)b_t(Z_t) + \mathcal{B}_0 Z_t\Big)dt + dW_t \\ dX_t = \mathcal{B}_1 Z_t dt \\ \mu_{Z_t}(dz) = u_t(z)dz, \end{cases} \tag{1.1}$$

for a given random initial condition

$$Z_0 = (V_0, X_0) \sim \mu_0 = u^0(z)(dz). \tag{1.2}$$

Here W is a d-dimensional Brownian motion, $\mathcal{B}_0 \in \mathcal{M}^{d\times N}$ and $\mathcal{B}_1 \in \mathcal{M}^{(N-d)\times N}$ are constant matrices, and $F : [0, +\infty) \to \mathcal{M}^{d\times m}$ is a matrix-valued Borel function. In order for the process (V, X) to take values on $\mathbb{R}^N$ one can formally consider $b : [0, T] \times \mathbb{R}^N \to \mathbb{R}^m$ in a way that the product in (1.1) can be made sense of as a usual multiplication *row-by-column*. One of the novel features here, however, is that $b_t(\cdot)$ is a distribution, hence the *product* $F(u)b$ needs to be defined with care. Another trait of (1.1) is that the unknown Z has a density u which appears in the drift of the V-component, and this makes the problem a McKean-Vlasov type SDE. Notice moreover that the SDE is degenerate, in the sense that the noise W appears only in the dynamics of the V-component. In order for the SDE to be well-posed, we will then require the hypoellipticity of the Kolmogorov operator $\mathcal{K}$ associated to (1.1) when $F \equiv 0$ (see Assumption 1.1).

We point out that, when $N = 2d$, the system (1.1) reduces to the *singular kinetic* McKean-Vlasov SDE

$$\begin{cases} dV_t = F\big(u_t(V_t, X_t)\big)b_t(V_t, X_t)dt + dW_t, \\ dX_t = V_t dt, \\ \mu_{(V_t,X_t)}(dvdx) = u_t(v, x)dvdx, \end{cases} \tag{1.3}$$

where the $\mathbb{R}^{2d}$-valued pair (V_t, X_t) describes *velocity* and *position* of a particle in the phase-space. When $F \equiv 1$ and b_t is a linear function of V_t (in particular there is no McKean-Vlasov interaction and b_t is not singular), (1.3) reduces to the classical Langevin model [45], which serves as a pilot example of more complex kinetic models (see [37], [36] among others). Recently, kinetic McKean-Vlasov models attracted the interest of several authors (e.g. [13, 33, 61, 63]). As it is pointed out in [33],(1.3) reflects the macroscopic behavior of a system of particles obeying Newton's second law with Gaussian noise, interacting through a mean-field force coefficient, with the singular nature of the latter accounting for environmental noise that acts on all particles. In [13] the authors study a conditional version of (1.3), as a Langrangian stochastic model alternative to Navier-Stokes equations for the simulation of turbulent flows. In [63] (see also [61]), the author studies a non-singular version of (1.3) where the McKean-Vlasov interaction also affects the diffusion term, motivated by the fact that the non-linear Fokker-Planck PDE associated to this class includes to the kinetic Landau equation as a prototype example, the latter being the most popular mathematical model in

the theory of collisional plasma. In this regards, the 2nd order part of the Fokker-Planck PDE associated to (1.3), which reads as

$$(\partial_t + \langle v, \nabla_x \rangle) u_t + \mathrm{div}_v\big(u_t F(u_t) b_t\big) = \Delta_v u_t, \tag{1.4}$$

differs from the one in the Landau equation, in particular it is linear. However, it is argued in [1] that a satisfactory description of a plasma is obtained when one also takes into account not only binary collisions (which are missing in (1.4)) but also collective effects modelled by a mean-field force term, which is present in (1.3) and reflected by the non-linear 1st order term in (1.4).

The main aim of this paper is to give a mathematical meaning to equation (1.1) and to prove existence and uniqueness of a solution, together with some regularity results on its time-marginal probability densities (see Theorem 1.6 and its proof in Section 5.2). To define the notion of solution to the McKean-Vlasov equation, which is a non-linear equation in the sense that the drift of the SDE depends on the solution via the density of the solution itself, a natural step is to define and investigate its linear counterpart, which is a kinetic-type SDE where the drift is singular but does not depend on the law of the solution itself. Such SDE is formulated as a martingale problem in Section 5.1, and results on its well-posedness are the first by-product of this work. We denote by $\mathcal{K}$ the parabolic generator of the underlying Gaussian process solving the SDE (1.1), see (2.11). Another by-product is the study of the nonlinear Fokker-Planck (FP) PDE given by

$$\begin{cases} \mathcal{K}' u_t = \mathrm{div}_v\big(u_t F(u_t) b_t\big), & t \in (0, T], \\ u_0 = u^0, \end{cases} \tag{1.5}$$

where $\mathcal{K}'$ is the formal adjoint of $\mathcal{K}$, given in (2.13)-(2.14), u^0 is the law density of $\mu_{(V_0, X_0)}$ and div_v denotes the divergence in the variable v of the vector (v, x). Indeed, we prove that the density of the solution to (1.1) does satisfy (1.5) and for this, analytical results for the latter are crucial (see Section 3), which might have separate interest. These include the existence and uniqueness of a weak solution in an anisotropic Hölder space, as well as stability results for the regularized version of (1.5), obtained by replacing the distribution b_t with a smooth function. Finally, we mention a third by-product, namely analytical results on the singular kinetic-type backward Kolmogorov PDE

$$\begin{cases} \mathcal{K} u + \langle \mathcal{B}, \nabla_v \rangle u = \lambda u + g, & t \in (0, T), \\ u_T = \phi, \end{cases} \tag{1.6}$$

with $\mathcal{K}$ being defined in (1.9), and $\mathcal{B}$ being a suitable distribution, which is the fundamental tool used to set up the linear martingale problem. In Section 4 we prove results that are analogous to those for the Fokker-Planck PDE (1.5). We also note that all the PDE results of this work rely on suitable Schauder estimates in anisotropic Besov spaces for the semigroups of $\mathcal{K}$ and $\mathcal{K}'$, which are stated in Theorem 2.10 and proved in Appendix B.

1.2 Literature review

As we have mentioned above, the system (1.1) that we will investigate combines three features: distributional coefficients (hence the term *singular*), the fact that the drift depends on the density of the law of the unknown itself (hence the term *McKean-Vlasov*), the fact that the Brownian motion drives only the V-component (hence the term *degenerate* or *kinetic*). All these features have previously been studied in the literature, as

we see below, but this is, to the best of our knowledge, the first time they are combined together when the drift depends in a pointwise fashion on the density law.

The first examples of SDEs with distributional drifts date back some 20 years ago with the works [7, 28, 29, 57] in one dimension. Some years later, [27] treat the case of any dimension $d \geq 1$ and with a drift in a fractional Sobolev space of negative regularity index $-\beta > -1/2$, in the so-called Young regime. Immediately after and independently [23] treat the case of singular coefficients in Besov spaces with negative regularity index $-\beta > -2/3$ in dimension 1 (i.e beyond the Young regime). In the following years [15] investigate extensions in dimension d and $-\beta > -2/3$. Another generalization is given simultaneously in [2] and [19], where the authors consider the extension to α-stable noises in place of Brownian motion, in the Young regime and dimension d, producing also an explicit description of the dynamics of the solution using a suitable semigroup. In [39] they frame the problem as a martingale problem and derive a description of the dynamics in terms of weak Dirichlet processes. One of the common tools that can be found in many of these works is a singular Kolmogorov equation like (1.6) in the uniformly parabolic case, i.e. when $N = d$, which is needed to cast the singular SDE as a martingale problem. A solution for the latter is a process Z and a probability $\mathbb{P}$ such that for all $f \in \mathcal{D}$ we have $f(t, Z_t) - f(0, X_0) - \int_0^t \mathcal{L}f(s, X_s)ds$ is a $\mathbb{P}$-martingale, where $\mathcal{L}$ is the 'generator' of Z (formally given by $\mathcal{L} = \mathcal{K} + \langle \mathcal{B}, \nabla_v \rangle$). The class of functions $\mathcal{D}$ is constructed as the set of solutions to (1.6) for smooth forcing terms g and terminal conditions ϕ. Once the martingale problem is defined, one uses results on the Kolmogorov PDE (1.6) (such as well-posedness and continuity with respect to the coefficients) combined with classical tools of stochastic analysis such as tightness and uniqueness results of marginal laws to show existence and uniqueness of a martingale solution. After these early foundational works on SDEs with singular coefficients, people started to explore this topic further: by studying numerical schemes for singular SDEs [22, 31]; allowing for different type of noises [43]; studying heat kernel estimates for singular SDEs [55]; studying *singular kinetic* SDEs, *singular McKean-Vlasov* SDEs and even *singular kinetic McKean-Vlasov* SDE, as we see below.

McKean-Vlasov equations were originally introduced by McKean in the mid 1960s and have become very popular, in recent years, after the well-known Saint-Flour lecture notes by A.S. Sznitman [59], and they typically arise as limit of interacting particle systems when the number of particles goes to infinity and when the dynamics of each particle depends on the empirical measure of all other particles. Depending on the type of interaction, one obtains different limiting equations. In the so-called moderate interaction, first introduced by [51] for smooth drifts, the particles interact through a mollified version of the empirical measure, and the limiting equation turns out to depend on the density of the law in a pointwise way (rather than the law itself). A similar McKean SDE whose marginal laws solve the porous media PDE was provided in [9]. Other McKean SDEs with more general diffusion and drift coefficients were investigated by [11, 5, 8, 4]. In the framework of moderate interaction, but with singular drift, and in the special non-degenerate case, the well-posedness of the limiting *singular McKean-Vlasov* SDE (1.1) has been recently studied in [40]. More precisely, in [40] the authors prove weak existence and uniqueness to the equation (1.1) in the special case when $\mathcal{B}_0 = 0$ and $d = N$, that is when the vector (V, X) in (1.1) reduces to the first component V, and for $-\beta > -1/2$. Similar equations have been studied in [16, 17] where they analyse a McKean-Vlasov type SDE with singular coefficients with a drift of convolution form $b * \mu_{V_t}$ and a general α-stable process as noise. They prove well-posedness of said SDE and heat kernel estimates. Notice that their threshold for the singularity is $-\beta > -1$ but they are still in the Young regime due to the smoothing effect of the convolution, as opposed to the pointwise product in (1.1), which is the same as in [40].

In the realm of *kinetic*-type SDEs, the well-posedness of the *linear* martingale problem, together with Gaussian estimates for the density, have been established in [50] and [49], assuming the coefficients in suitable Hölder spaces induced by an anisotropic norm that takes into account the different time-scaling properties of the degenerate components, namely the component X of (V, X) according to our notation. In the latter references, the drift of the degenerate component X may be non-linear and the diffusion coefficient is allowed to be non-constant, provided that it satisfies a uniform ellipticity condition on $\mathbb{R}^d$. In [18], the authors allow the drift of V to be an unbounded discontinuous function and derive heat kernel and gradient kernel estimates. In [33], the drift of V is a convolution with a kernel that may be a distribution and the authors study the singular martingale problem in order to tackle a convolutional kinetic McKean-Vlasov SDE (see also the next paragraph). In the *kinetic McKean-Vlasov* setting, we mention a few contributions. In [13] a conditional Langevin-type dynamics is considered as an alternative approach to Navier-Stokes equations for turbulent flows. The weak well-posedness of the McKean-Vlasov SDE was proved via propagation of chaos, by considering a smoothed SDE and proving that the corresponding interacting particle system propagates chaos. The weak existence for equations with McKean-Vlasov coefficients in the form of $\mathbb{E}[b(t, z, Z_t)]$, both in the drift and the diffusion, was recently proved in [61] for the strictly kinetic case ($N = 2d$), under minimal regularity conditions (uniform continuity w.r.t. to the position variable in both the state and the measure component) on the function b. Similar results were recently proved in [63].

Finally, we mention [33] in which the authors combine all three aspects considered in this paper: *singular, kinetic* and *McKean-Vlasov*. In particular, they study a kinetic McKean-Vlasov equation which differs from (1.3) in that the dynamics of the velocity component V_t reads as

$$dV_t = b_t(Z_t)dt + (K * \mu_{X_t})(X_t)dt + dW_t, \tag{1.7}$$

where the drift component is the sum of a term b_t in a weighted anisotropic Besov space with regularity index $-\beta$ with $\beta \in (1/2, 2/3)$ (hence this is outside the Young regime and requires the use of tools like paracontrolled calculus) and one convolution part involving the law of the position X with a singular kernel K with regularity index $\alpha > (\beta - 1)/3$ (notice that α can be negative). They prove existence of a martingale solution, and under the stronger assumptions that K is bounded and measurable (thus non-singular) they also prove uniqueness. One of the main differences between the setting in [33] and that of the present work regards the way the drift depends on the law: (i) in (1.7) the dependence appears through the convolution with a singular term, hence it is linear, while here, as in (1.1), we have a nonlinear function evaluated pointwisely in the law and multiplied by a Schwartz distribution; (ii) in (1.7) the dependence on the law, and thus the mean-field interaction of the associated particle system, only involves the position variable X.

Finally we stress that in the present work we derive our results in the general degenerate setting of equation (1.1) (with $0 < d < N$) and not only in the purely kinetic case ($N = 2d$).

As mentioned in Section 1.1, a key step to prove our results on the SDE (1.7) is the study of the singular non-linear Fokker-Planck equation (1.5) and the backward Kolmogorov equation (1.6). From the PDE point of view, our results in Sections 3 and 4 on the well-posedness, regularity and stability of these equations are natural extensions of those in [40] and [38] from the elliptic to the hypoelliptic setting. As the proofs in the latter references rely on Besov-Schauder estimates for the heat semigroup (see [32]), our proofs rely on their anisotropic counterpart for the semigroup of the operators $\mathcal{K}$ and $\mathcal{K}'$ (defined in (1.9) and (2.13)), for which we provide a complete proof. To the best of our knowledge, in the hypoelliptic framework, these estimates only appeared in [64] for the strictly kinetic ($N = 2d$) and homogeneous case, namely when the

operator $\mathcal{K}$ is invariant with respect to a family of dilations. We also mention that our PDE results are somehow complementary to those in [33], where the paracontrolled case ($\beta \in (1/2, 2/3)$) was considered, in the strictly kinetic framework. In general, there is a large body of literature, also very recent, concerned with the regularity properties of degenerate-type PDEs involving $\mathcal{K}$ or its adjoint $\mathcal{K}'$ as a principal operator. In the Hölder setting, we recognize two main approaches. In the semigroup approach, initiated by [48], solutions are defined in the distributional sense: in this framework, solutions do not benefit from the time-smoothing effect that is typical of parabolic equations, and the regularity structures are typically defined in terms of spatial *anisotropic* Hölder spaces (see Section 2.3), which do not require Hölder-regularity with respect to the time-variable. With respect to this approach, the anisotropic Besov spaces defined in Section 2.1, and the related Schauder estimates for the semigroups of $\mathcal{K}$ and $\mathcal{K}'$ in Section 2.2, are the natural distributional extensions. On the other hand, in the stream of research started by [56], solutions in the Lie sense are defined in terms of directional derivatives along the Hörmander vector fields. In this approach, the regularity of the coefficients and of the solutions is typically stated in so-called *instrinsic* Hölder spaces (see Section 2.3), where the regularity properties in space and time are strictly intertwined. Some authors (see [10], [25], [47]) have recently combined these two approaches to obtain intrinsic regularity of the solution (with respect to all the variables) by only assuming anisotropic regularity for the data (and only measurability). We refer to [46] and [47] for a more exhaustive overview of the literature regarding these two approaches and their combination. In particular, we refer to [46] and to the previous work [24] for the study of fundamental solutions in the degenerate Hölder setting. We note that, although the theory of functional intrinsic spaces has already been developed in the Sobolev setting (e.g. [30], [54]), the distributional Besov counterpart of the intrinsic Hölder spaces has not yet been studied and it is an open problem for future research. This will be necessary to establish sharp regularity estimates for the singular PDEs considered in this paper.

1.3 Setting and main result

Throughout the paper, we set the matrix

$$B := \begin{pmatrix} \mathcal{B}_0 \\ \mathcal{B}_1 \end{pmatrix} \in \mathcal{M}^{N\times N}, \tag{1.8}$$

where $\mathcal{B}_0 \in \mathcal{M}^{d\times N}$ and $\mathcal{B}_1 \in \mathcal{M}^{(N-d)\times N}$ are the real matrices in (1.7). We also denote by Δ_v and ∇_z, respectively, the Laplacian operator on $\mathbb{R}^d$ and the gradient on $\mathbb{R}^N$, namely

$$\Delta_v = \sum_{i=1}^{d} \partial^2_{v_i}, \qquad \nabla_z = (\partial_{z_1}, \cdots, \partial_{z_N}).$$

Assumption 1.1. The Kolmogorov operator

$$\mathcal{K} := \mathcal{A} + \partial_t := \frac{1}{2}\Delta_v + \underbrace{\langle Bz, \nabla_z\rangle + \partial_t}_{=:Y}, \qquad z = (v, x) \in \mathbb{R}^N, \tag{1.9}$$

is hypoelliptic on $\mathbb{R}^{N+1}$.

Remark 1.2. By Hörmander theorem, $\mathcal{K}$ in (1.9) is hypoelliptic if and only if the vector fields $\partial_{v_1}, \ldots, \partial_{v_d}$ and Y satisfy

$$\text{rank Lie}(\partial_{v_1}, \ldots, \partial_{v_d}, Y) = N + 1. \tag{1.10}$$

Example 1.3. The kinetic equation (1.3) can be obtained from (1.7) by setting $N = 2d$, $m = d$ and

$$B := \begin{pmatrix} 0 & 0 \\ I_d & 0 \end{pmatrix}.$$

In this case we have $Y = \partial_t + \langle v, \nabla_x \rangle$ and

$$[\partial_{v_i}, Y] = \partial_{x_i}, \qquad i = 1, \cdots, d,$$

thus (1.10) is satisfied.

In [44] it was shown that (1.10), known as weak (or parabolic) *Hörmander* condition, is equivalent to $\mathcal{B}_1$ having the block-form

$$\mathcal{B}_1 = \begin{pmatrix} B_1 & * & \cdots & * & * \\ 0 & B_2 & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & B_r & * \end{pmatrix} \tag{1.11}$$

where the $*$-blocks are arbitrary and B_j is a $(d_{j-1} \times d_j)$-matrix of rank d_j with

$$d \equiv d_0 \geq d_1 \geq \cdots \geq d_r \geq 1, \qquad \sum_{i=0}^{r} d_i = N. \tag{1.12}$$

It is convenient to set

$$\bar{d}_i := \sum_{k=0}^{i} d_k, \qquad i = 0, \cdots, r, \qquad \text{and} \quad \bar{d}_{-1} := 0.$$

Next we introduce the notation for function spaces used in this paper. For a given $T > 0$, $\alpha \in \mathbb{R}$, and for any $(U, \|\cdot\|_U)$ generic normed space endowed with the Borel σ-algebra generated by the topology induced by $\|\cdot\|_U$, we introduce the following notations.

- $\mathcal{S}$ the space of Schwartz functions and $\mathcal{S}'$ the space of tempered distribution on $\mathbb{R}^N$.
- $\mathcal{C}_b$ the space of bounded continuous functions on $\mathbb{R}^N$, equipped with the norm $\|\cdot\|_{L^\infty} = \|\cdot\|_{L^\infty(\mathbb{R}^N)}$.
- $\mathcal{C}_c^\infty$ the space of $\mathcal{C}^\infty$ functions on $\mathbb{R}^N$ with compact support.
- $\mathcal{C}_B^\alpha$ the anisotropic Besov space of order α and by $\|\cdot\|_\alpha$ its norm, see Definition 2.6.
- $C_{B,T}^\alpha$ the space of intrinsically α-Hölder continuous functions for $\alpha > 0$, and by $\|\cdot\|_{C_{B,T}^\alpha}$ its norm, see Definition 2.15.
- $L_T^\infty U$, the space of Borel-measurable functions $f : [0, T] \to U$ such that

$$\|f\|_{T,U} := \operatorname*{ess\,sup}_{t \in [0,T]} \|f_t\|_U < +\infty.$$

  We have two special cases:

  – when $U = \mathcal{C}_B^\alpha$ the norm $\|f\|_{T,U}$ is denoted by $\|f\|_{T,\alpha}$,

  – when $U = \mathcal{C}_b$, the norm $\|f\|_{T,U}$ is denoted by $\|f\|_T$.

- $C_T U$ the space of continuous functions $f : [0,T] \to U$.
- $C_T \mathcal{S}$ the space of continuous functions $f : [0,T] \to \mathcal{S}$, where $\mathcal{S}$ is equipped with the usual strong topology.

We use the notation $\langle \cdot, \cdot \rangle$ to indicate the scalar product in $\mathbb{R}^N$, and $\langle \cdot | \cdot \rangle$ for the distributional dual pairing in $\mathcal{S}$ and $\mathcal{S}'$.

We now specify the regularity assumptions on the coefficients b, F in (1.1) and on the density of the initial condition u^0 in (1.2), under which we will prove the main result.

Assumption 1.4. Let $T > 0$, $\beta \in (0, \frac{1}{2})$. We have

(i) $b \in C_T \mathcal{C}_B^{-\beta}$.

(ii) $u^0 \in \mathcal{C}_B^{\beta+\varepsilon}$ for some $\varepsilon \in (0, 1-2\beta)$, and is a probability density on $\mathbb{R}^N$.

Let $F : \mathbb{R} \to \mathcal{M}^{d\times m}$ and $\widetilde{F} : \mathbb{R} \to \mathcal{M}^{d\times m}$, defined as

$$\widetilde{F}(s) := sF(s). \tag{1.13}$$

We will make use of the following assumptions for $\Phi = F$ or $\Phi = \widetilde{F}$.

Assumption 1.5. The function $\Phi : \mathbb{R} \to \mathcal{M}^{d\times m}$ is differentiable, with bounded and globally Lipschitz-continuous derivative.

Postponing until Section 5.2 (Definition 5.14) the precise definition of solution (in law) to (1.1) (denoted as solution to MKMP(F, b, u^0)), we formulate the following main result, which will be restated and proved in Theorem 5.15. Concerning the regularity of the density of the solution to the MKMP(F, b, u^0), we refer to Remark 5.16.

Theorem 1.6. *Let Assumptions 1.1 and 1.4 hold. Let also Assumption 1.5 hold with $\Phi = F, \widetilde{F}$. Then there exists a unique solution to MKMP(F, b, u^0).*

In order to help the reader better understand the paper, we now explain the key steps and overall strategy for the proof of the main result. The idea in a nutshell is to construct a reasonably good "guess" of solution to the McKean problem and then verify that such a guess is the unique solution of our problem. However, the construction of the guess solution is not trivial.

The *first step*, in order to do this, is to view the MKMP(F, b, u^0) as a particular case in a class of *linear* martingale problems, denoted by MP$(\mathcal{B}, \mu_0)$, where the drift $\mathcal{B}$ does not contain any unknowns. Here $\mathcal{B}$ must be any given Besov distribution and μ_0 any initial law. We will construct our guess solution by solving the linear martingale problem, which is an easier environment to work in. It can be easily verified that a solution of MKMP(F, b, u^0) is a solution of MP$(F(u)b, u^0(z)dz)$, so in order to move from the McKean-Vlasov framework to the linear framework we need to know the density of the solution itself, because the unknown density u of (1.1) has to be given as an input of the linear problem. In the linear case, by applying the Itô formula, it is easily verified that the marginals of an SDE are (distributional) solutions of the associated FP equation. This fact also holds for a solution of MKPM(F, b, u^0), as proved in Proposition 5.13, even though the associated FP equation is non-linear in this case. The results on well-posedness for the nonlinear FP equation are proved in Section 3, and are a first hint about the well-posedness of (1.1). Once we are in the

linear framework, the *second step* consists of finding a (unique) solution to the singular linear martingale problem, which is the content of Theorem 5.11. To do so, the singular drift is smoothed and the result [26, Theorem 2.6] allows us to find the solution of the Stroock-Varadhan martingale problem. At this point, it is worth noticing that in the smoothed case our notion of solution coincides with the classical one à là Stroock-Varadhan, as proven in Proposition 5.5. To show *existence* of a solution to the non-smoothed linear problem we have to verify that the sequence of measures that solves the regularized problem is tight, and that the limit of a subsequence is a solution. To prove tightness (see Lemma 5.7) we make use of a uniform-boundness result of solutions to a Kolmogorov-type PDE, whose analysis is done in Section 4. To prove *uniqueness*, we extend a well-known uniqueness propriety of martingale problems to time-inhomogeneous coefficients, which is referred to as 'Propriety P' in this work (see Appendix A).

The *final step* is now easy and consists of checking that the solution constructed as described above is indeed the unique solution to the singular McKean-Vlasov martingale problem. This can be seen using tightness as well as continuity results of the Zvonkin-type PDE and the Fokker-Planck PDE, together with the fact that for the smoothed coefficient the result is valid. This is done in Section 5.2.

The rest of the paper is organized as follows. In Section 2 we introduce the anisotropic Besov and Hölder spaces we will work in and the main analytical tools that are needed to study the non-linear Fokker-Planck and the backward Kolmogorov PDEs. In particular we introduce the semigroups associated to the Kolmogorov backward operator and to the Fokker-Planck forward operator and provide key Schauder estimates for both, the proof of which is postponed to the Appendix. In Section 3.1 we study the well posedness of the nonlinear Fokker-Planck equation via local fixed point arguments and using Schauder estimates. We also show continuity results for the FP equation in Section 3.2. Section 4 is devoted to showing well-posedness and continuity results for the linear Kolmogorov equation. Section 5.1 contains the study of the linear martingale problem and Section 5.2 the well posedness of the McKean-Vlasov problem. Appendix A contains the extension of the uniqueness results for time-inhomogeneous martingale problems knowing the uniqueness of the marginals and Appendix B the proof of Schauder estimates.

2 Analytical tools

In this section we provide the main analytical results that are needed for the study of the singular non-linear Fokker-Planck PDE (1.5) and backward Kolmogorov PDE (1.6). We first introduce the relevant anisotropic Besov spaces and recall some of their basic properties (Section 2.1), then we state the inherent Schauder estimates for the semigroup kernels of $\mathcal{K}$ and $\mathcal{K}'$ (Section 2.2). We then recall the definition of anisotropic Hölder spaces, their characterization as Besov spaces with non-integer positive regularity index, and briefly compare them with their intrinsic counterparts (Section 2.3). Finally we state some useful regularization results (Section 2.4).

Hereafter we will always assume that Assumption 1.1 is satisfied and thus $\mathcal{B}_1$ admits a representation as in (1.11)-(1.12), which factorizes $\mathbb{R}^N$ in $r+1$ components of dimensions $d_0 \geq \cdots \geq d_r \geq 1$. Accordingly, we will employ the notation $z = (z_0, \cdots, z_r) \in \mathbb{R}^N$ with each $z_i \in \mathbb{R}^{d_i}$.

2.1 Anisotropic Besov Spaces

For the reader's convenience, we recall the classical Fourier-based construction of anisotropic distributional Besov spaces together with some useful results. We focus on the particular case of integrability and microscopic parameters equal to infinity (typically denoted by p and q in the literature), as this work only covers this case. The interested reader can refer to [60, Chapter 5] for the general definition, and to the references therein for an account on the historical development of these spaces.

Let us define the so-called *anisotropic norm* as

$$|z|_B := \sum_{i=0}^{r} |z_i|^{\frac{1}{2i+1}}, \qquad z = (z_0, \cdots, z_r) \in \mathbb{R}^N,$$

where $|z_i|$ is the Euclidean norm in $\mathbb{R}^{d_i}$. It is immediate that $|\cdot|_B$ is not strictly a norm as it is not homogeneous with respect to usual scalar multiplication. However, homogeneity holds for the family of dilations $D_\lambda : \mathbb{R}^N \longrightarrow \mathbb{R}^N$, $\lambda > 0$, defined as

$$D_\lambda z = \lambda.z := (\lambda z_0, \lambda^3 z_1, \ldots, \lambda^{2r+1} z_r), \qquad z = (z_0, \cdots, z_r) \in \mathbb{R}^N. \tag{2.1}$$

In this work $\mathbf{B}_\tau(z)$ denotes the closed anisotropic disk, or ball, centred at $z \in \mathbb{R}^N$ with radius $\tau > 0$, i.e. $\mathbf{B}_\tau(z) := \{y \in \mathbb{R}^N : |y - z|_B \leq \tau\}$. The ball centred at the origin is denoted by $\mathbf{B}_\tau := \mathbf{B}_\tau(0)$. Moreover, we define the anisotropic annuli

$$\mathcal{A}_{\tau,\rho} := \{x \in \mathbb{R}^N : \tau \leq |x|_B \leq \rho\}, \qquad 0 < \tau < \rho.$$

We finally denote by Q the so-called *homogeneous dimension*, set as

$$Q := \sum_{i=0}^{r} d_i(2i+1).$$

As in the isotropic case, the anisotropic spaces are defined through a Paley-Littlewood decomposition. We start by considering a smooth, non-negative function ρ_{-1} on $\mathbb{R}^N$ such that

$$\rho_{-1} \equiv 1 \text{ on } \mathbf{B}_{\frac{1}{2}}, \qquad \text{supp}\rho_{-1} \subseteq \mathbf{B}_{\frac{2}{3}} \quad \text{and} \quad \int_{\mathbb{R}^N} \rho_{-1}(\xi) d\xi = 1. \tag{2.2}$$

The sequence of functions $\{\rho_j\}_{j\geq 0}$, defined by

$$\rho_j(\xi) := \rho_{-1}(2^{-(j+1)}.\xi) - \rho_{-1}(2^{-j}.\xi), \qquad \xi \in \mathbb{R}^N,$$

is called an anisotropic *Paley-Littlewood partition*.

Remark 2.1. By definition we have the following properties:

$$\begin{aligned}
&(a) \quad \rho_j(\xi) = \rho_0(2^{-j}.\xi), \qquad j \geq 0,\\
&(b) \quad \text{supp}\rho_j \subseteq \mathcal{A}_{2^{j-1},2^{j+1}}, \qquad j \geq 0,\\
&(c) \quad \text{supp}\rho_i \cap \text{supp}\rho_j \neq \emptyset \Longleftrightarrow |i-j| \leq 1, \qquad i,j \geq -1,\\
&(d) \quad \sum_{j=-1}^{n} \rho_j(\xi) = \rho_{-1}\big(2^{-(n+1)}.\xi\big), \qquad n \geq -1,\\
&(e) \quad \sum_{j=-1}^{+\infty} \rho_j \equiv 1 \text{ on } \mathbb{R}^N.
\end{aligned} \tag{2.3}$$

For any $f \in L^1(\mathbb{R}^N)$ we employ the following definitions for the Fourier transform and its inverse:

$$\mathcal{F}(f)(\xi) = \widehat{f}(\xi) := (2\pi)^{-\frac{N}{2}} \int_{\mathbb{R}^N} e^{-i\langle z,\xi\rangle} f(z)dz,$$
$$\mathcal{F}^{-1}(f)(z) = \check{f}(z) := (2\pi)^{-\frac{N}{2}} \int_{\mathbb{R}^N} e^{i\langle z,\xi\rangle} f(\xi)d\xi.$$

Remark 2.2. It can be directly verified that the scaling property of the Fourier transform for the $\{\rho_j\}_{j\geq 0}$ reads as

$$2^{jQ}\check{\rho}_0 \circ D_{2^j} = \mathcal{F}^{-1}(\rho_0 \circ D_{2^{-j}}) = \check{\rho}_j, \qquad j \geq 0. \tag{2.4}$$

Another important property is

$$(2\pi)^{-\frac{N}{2}} \int_{\mathbb{R}^N} \check{\rho}_{-1}(z)dz = 1, \qquad (2\pi)^{-\frac{N}{2}} \int_{\mathbb{R}^N} \check{\rho}_j(z)dz = 0, \quad j \geq 0,$$

which stems from

$$(2\pi)^{-\frac{N}{2}} \int_{\mathbb{R}^N} \check{\rho}_j(z)dz = \mathcal{F}\left(\mathcal{F}^{-1}\rho_j\right)(0) = \rho_j(0).$$

The Fourier transform of $f \in \mathcal{S}'$ is defined, as usual, by duality, i.e.

$$\langle \mathcal{F}(f) | \varphi \rangle = \langle \widehat{f} | \varphi \rangle := \langle f | \widehat{\varphi} \rangle, \qquad \varphi \in \mathcal{S},$$

and the same for $\mathcal{F}^{-1} = \check{f}$. For $f \in \mathcal{S}'$, the j-th Paley-Littlewood block of f is defined as

$$\Delta_j f := \mathcal{F}^{-1}\big(\rho_j \widehat{f}\,\big), \qquad j \geq -1. \tag{2.5}$$

Remark 2.3. By the properties of the Fourier transform, we have

$$\Delta_j f = \check{\rho}_j * f,$$

and in particular, $\Delta_j f$ is smooth. On the other hand, it is easy to check that $\sum_{j=-1}^{n} \Delta_j f$ tends to f in $\mathcal{S}'$ as $n \to \infty$.

The following result, whose proof is completely analogous to that of [3, Lemma 2.1] in the isotropic case, is fundamental to comprehend the construction of Besov spaces used in the paper.

Theorem 2.4 (Bernstein's type inequality)**.** *For every $k \in \mathbb{N}_0$, there exists a positive constant $C = C(k)$ such that*

$$\sum_{l=\bar{d}_{i-1}+1}^{\bar{d}_i} \|\partial_{z_l}^k f\|_{L^\infty} \leq C\lambda^{(2i+1)k}\|f\|_{L^\infty}, \qquad \lambda > 0, \quad i = 0, \cdots, r, \tag{2.6}$$

for any $f \in L^\infty(\mathbb{R}^N)$ such that $\operatorname{supp}\widehat{f} \subseteq \mathbf{B}_\lambda$. *Note that, in particular, C does not depend on λ.*

By applying this result to the Paley-Littlewood partition, we have the following.

Corollary 2.5. *For any $f \in \mathcal{S}'$ and $k \in \mathbb{N}_0$, we have*

$$\sum_{l=\bar{d}_{i-1}+1}^{\bar{d}_i} \|\Delta_j(\partial_{z_l}^k f)\|_{L^\infty} \leq C\, 2^{j(2i+1)k}\|\Delta_j f\|_{L^\infty}, \qquad i = 0, \cdots, r, \qquad j \geq -1, \tag{2.7}$$

where $C = C(k) > 0$.

Proof. By (2.5) we have

$$\mathcal{F}(\Delta_j f) = \rho_j \widehat{f}, \qquad j \geq -1,$$

whose support is contained in $\mathbf{B}_{2^{j+1}}$ by (2.3). If $\|\Delta_j f\|_{L^\infty} = \infty$, there is nothing to prove. Otherwise, since the decomposition operators Δ_j commute with the derivatives, the result follows from (2.6). □

We can now give the definition of anisotropic Besov space.

Definition 2.6. Given $\gamma \in \mathbb{R}$, the space $\mathcal{C}_B^\gamma$ of the distributions $f \in \mathcal{S}'$ such that

$$\|f\|_\gamma := \sup_{j \geq -1} \left(2^{j\gamma} \|\Delta_j f\|_{L^\infty}\right) < \infty$$

is called *anisotropic Besov space of order* γ. In this context, γ is called the *regularity parameter*.

It is straightforward to see that

$$\alpha < \gamma \Longrightarrow \mathcal{C}_B^\gamma \subset \mathcal{C}_B^\alpha.$$

Also, by applying inequality (2.7), for any $\gamma \in \mathbb{R}$ and $k \in \mathbb{N}_0$ we have

$$\sum_{l=\bar{d}_{i-1}+1}^{\bar{d}_i} \|\partial_{z_l}^k f\|_\gamma \leq C \|f\|_{\beta+(2i+1)k}, \qquad i = 0, \cdots, r. \tag{2.8}$$

The quantity $(2i+1)$ represents the reduction in terms of regularity whenever a derivative along the i-th block is required.

Example 2.7. We have that $L^\infty(\mathbb{R}^N) \subset \mathcal{C}_B^0$. Indeed, for any $f \in L^\infty(\mathbb{R}^N)$, Young's inequality yields

$$\|\Delta_j f\|_{L^\infty} \leq \|\check{\rho}_j\|_{L^1} \|f\|_{L^\infty}.$$

Also, by (2.4) we have $\|\check{\rho}_j\|_{L^1} = \|\check{\rho}_0\|_{L^1}$, and thus

$$\|f\|_0 = \sup_{j \geq -1} \|\Delta_j f\|_{L^\infty} < \infty.$$

Example 2.8. The Dirac delta δ_0 of $\mathbb{R}^N$ is an element of $\mathcal{C}_B^\gamma$ if and only if $\gamma \leq -Q$. To see this, note that

$$\Delta_j \delta_0 = \check{\rho}_j * \delta_0 = \check{\rho}_j, \qquad j \geq -1.$$

Therefore, by (2.4) we obtain

$$\sup_{j \geq 0} \left(2^{j\gamma} \|\Delta_j \delta_0\|_{L^\infty}\right) = \sup_{j \geq 0} \left(2^{(j(\gamma+Q)} \|\check{\rho}_0\|_{L^\infty}\right),$$

which is finite if and only if $\gamma \leq -Q$.

To make sense of the product appearing in the MKV SDE (1.1) and in the PDEs (1.5) and (1.6), we will consider the so-called *Bony's product* (see [12]) between Besov distributions, which extends the product between functions (which are also tempered distributions). Intuitively, this is defined if we can make sense of the formal product

$$fg = \sum_{j,i=-1}^{\infty} \Delta_j f \Delta_i g, \tag{2.9}$$

which is possible if the sum between the regularity parameters of f and g is strictly positive. As in the isotropic framework, this is achieved by bounding separately the *paraproduct* and the *resonant* components of (2.9) (see [32, Section 2.1]). In particular, the following result can be proved by proceeding exactly as in [32, Lemma 2.1] (see also [33]).

Proposition 2.9. *Let $\alpha, \gamma \in \mathbb{R}$ with $\alpha + \gamma > 0$. Then Bony's product restricted to $\mathcal{C}_B^\alpha \times \mathcal{C}_B^\gamma$ is a bilinear form from $\mathcal{C}_B^\alpha \times \mathcal{C}_B^\gamma$ to $\mathcal{C}_B^{\alpha\wedge\gamma}$ and there exists a positive constant C such that*

$$\|fg\|_{\alpha\wedge\gamma} \leq C\,\|f\|_\alpha \|g\|_\gamma, \qquad f \in \mathcal{C}_B^\alpha,\ g \in \mathcal{C}_B^\gamma. \tag{2.10}$$

2.2 Kinetic-type semigroups and Schauder's estimates

In this section we present Besov-Schauder's estimates for the fundamental solutions of the hypoelliptic operators $\mathcal{K}$ in (1.9) and of its formal adjoint $\mathcal{K}'$ defined below, which appear in (1.6) and (1.5), respectively. Note that $\mathcal{K}$ is the Kolmogorov operator associated to the $\mathbb{R}^N$-valued linear SDE

$$dZ_t = BZ_t dt + \sigma dW_t, \tag{2.11}$$

with

$$\sigma = \begin{pmatrix} I_d \\ 0_{N-d,d} \end{pmatrix} \in \mathcal{M}^{N\times d},$$

where I_d and $0_{N-d,d}$ denote, respectively, the $d \times d$ identity matrix and the $(N-d) \times d$ matrix with null entries. We recall that B is as in (1.8)-(1.11)-(1.12), so that Assumption 1.1 is satisfied. Note that (2.11) coincides with (1.1) upon posing the McKean-Vlasov coefficient $F\big(u_t(Z_t)\big)b_t(Z_t)$ equal to zero. It is known since [34] (see also [53, Section 9.5] for a recent presentation) that $\mathcal{K}$ admits a unique smooth fundamental solution $\Gamma(T-t; y, z)$, $t < T$, $y, z \in \mathbb{R}^N$. Namely, the function $(t, y) \mapsto \Gamma(T-t; y, z)$ solves $\mathcal{K}u = 0$ on $(-\infty, T) \times \mathbb{R}^N$ and

$$\Gamma(T-t; \eta, z)dz \longrightarrow \delta_y(dz) \quad \text{weakly as } (t, \eta) \to (T^-, y).$$

In particular, Γ is explicitly given by

$$\Gamma(s; y, z) = \Gamma_s(z - e^{sB}y), \qquad s > 0, \quad z, y \in \mathbb{R}^N,$$

with

$$\Gamma_s(z) := \frac{(2\pi)^{-\frac{N}{2}}}{\sqrt{\det \mathcal{C}(s)}} \exp\left(-\frac{1}{2}\langle \mathcal{C}^{-1}(s)z, z\rangle\right), \tag{2.12}$$

and

$$\mathcal{C}(s) := \int_0^s e^{\tau B} A \left(e^{\tau B}\right)^\top d\tau, \qquad A := \begin{pmatrix} I_d & 0 \\ 0 & 0 \end{pmatrix}.$$

Note that $\mathcal{C}(s)$ is strictly positive definite if (and only if) Assumption 1.1 holds. It is also known that $\Gamma(t-s; y, z)$, $t > s$, is the fundamental solution (with respect to the forward variables t, z) of the Fokker-Planck operator $\mathcal{K}'$ given by

$$\mathcal{K}' := -\partial_t + \mathcal{A}' \tag{2.13}$$

$$\mathcal{A}' := \frac{1}{2}\Delta_v - \langle \nabla_z, Bz\rangle = \frac{1}{2}\Delta_v - \langle Bz, \nabla_z\rangle - Tr(B), \qquad z = (v, x) \in \mathbb{R}^N. \tag{2.14}$$

This means that the function $(t, z) \mapsto \Gamma(t-s; y, z)$ solves $\mathcal{K}'u = 0$ on $(s, +\infty) \times \mathbb{R}^N$ and

$$\Gamma(t-s; y, \eta)dy \longrightarrow \delta_z(dy) \quad \text{weakly as } (t, \eta) \to (s^+, z). \tag{2.15}$$

To fix the ideas, this implies that the function

$$v(s; t, z) := \int_{\mathbb{R}^N} \Gamma(t-s; y, z)\varphi(y)dy - \int_s^t \int_{\mathbb{R}^N} \Gamma(t-\tau; y, z)g(s, y)dyd\tau, \qquad t > s,\ z \in \mathbb{R}^N,$$

is a smooth solution to the Cauchy problem

$$\begin{cases} \mathcal{K}' v(s;\cdot,\cdot) = g, & \text{on } (s,+\infty)\times\mathbb{R}^N, \\ v(s;s,\cdot) = \varphi, \end{cases}$$

for any $\varphi \in C_b(\mathbb{R}^N)$ and g sufficiently smooth. An analogous representation holds for the solution to the backward Cauchy problem for $\mathcal{K}$.

In light of this, we want to extend the action of the fundamental solution to tempered distributions. For $t>0$, we consider $P_t, P_t' : \mathcal{S} \to \mathcal{S}$ acting as

$$P_t\varphi(y) = \int_{\mathbb{R}^N} \Gamma(t;y,z)\varphi(z)dz, \qquad P_t'\varphi(z) = \int_{\mathbb{R}^N} \Gamma(t;y,z)\varphi(y)dy,$$

which define semigroups. By duality, we then extend these semigroups $(P_t)_{t>0}$ and $(P_t')_{t>0}$ of $\mathcal{K}$ and $\mathcal{K}'$, respectively, as the family of linear operators $P_t, P_t' : \mathcal{S}' \to \mathcal{S}'$ acting as

$$\langle P_t f \,|\, \varphi\rangle := \langle f \,|\, P_t'\varphi\rangle, \qquad \langle P_t' f \,|\, \varphi\rangle := \langle f \,|\, P_t\varphi\rangle, \qquad \varphi \in \mathcal{S}, \quad f \in \mathcal{S}'.$$

Theorem 2.10 (Schauder's estimates)**.** *For any $\gamma \in \mathbb{R}$, $\alpha \geq 0$ and $T>0$ there exists a positive constant $C = C(\alpha,\gamma,T)$ such that, for any $f \in \mathcal{C}_B^\gamma$, we have*

$$\|P_t f\|_{\gamma+\alpha} + \|P_t' f\|_{\gamma+\alpha} \leq C\, t^{-\frac{\alpha}{2}} \|f\|_\gamma, \qquad t \in (0,T]. \tag{2.16}$$

The proof of Theorem 2.10 is postponed to Appendix B. These estimates generalize those in [32] for the semigroup of the heat operator, and those in [64] for the strictly kinetic ($N=2d$) and homogeneous case, namely when the operator $\mathcal{K}$ is invariant with respect to the family of dilations in (2.1).

2.3 Anisotropic and intrinsic Hölder spaces

As mentioned in the introduction, the anisotropic Besov spaces defined above represent a natural extension of the anisotropic Hölder spaces typically used in the study of $\mathcal{K}$ and $\mathcal{K}'$-type operators according to the semi-group approach. For $\gamma>0$, a full characterization of $\mathcal{C}_B^\gamma$ in terms of an anisotropic Hölder space can be found in [64, Lemma 2.8]. To illustrate the idea, we provide a representation of such a characterization for some ranges of β.

Definition 2.11. We denote by C_B^γ the linear space of functions $f : \mathbb{R}^N \longrightarrow \mathbb{R}$ such that the following quantity is finite:

$$\|f\|_{C_B^\gamma} := \begin{cases} \|f\|_{L^\infty} + \sup\limits_{z\in\mathbb{R}^N} \sup\limits_{h\in\mathbb{R}^N} \frac{|f(z+h)-f(z)|}{|h|_B^\gamma}, & \gamma \in (0,1], \\ \|f\|_{L^\infty} + \sup\limits_{z\in\mathbb{R}^N} \sup\limits_{h\in\mathbb{R}^{N-d}} \frac{|f(z+(0,h))-f(z)|}{|(0,h)|_B^\gamma} + \sum\limits_{l=1}^{d} \|\partial_{z_l} f\|_{C_B^{\gamma-1}}, & \gamma \in (1,3], \\ \|f\|_{L^\infty} + \sup\limits_{z\in\mathbb{R}^N} \sup\limits_{h\in\mathbb{R}^{N-d-d_1}} \frac{|f(z+(0,0,h))-f(z)|}{|(0,0,h)|_B^\gamma} + \sum\limits_{l=1}^{d} \|\partial_{z_l} f\|_{C_B^{\gamma-1}} + \sum\limits_{l=1}^{d_1} \|\partial_{z_{d+l}} f\|_{C_B^{\gamma-3}}, & \gamma \in (3,4]. \end{cases}$$

Previous quantity is a norm and C_B^γ is a Banach space.

Note that, up to order $\gamma = 3$, the only derivatives appearing are with respect to the first d space-variables. The first-order derivatives with respect to the space variables in the d_1-block only appear when $\gamma > 3$. We refer the reader to [64, Subsection 2.1] for the definition of C_B^γ for an arbitrary $\gamma > 0$. The basic idea is that a derivative with respect to a space-variable within the j-th block is counted as a $(2j+1)$-th order derivative.

Remark 2.12. If $\alpha > \gamma > 0$ then the space C_B^α is topologically embedded in C_B^γ.

The relation between the Besov spaces $\mathcal{C}_B^\gamma$ and C_B^γ is as follows (see [64, Lemma 2.8], which is an adaptation of [21, Proposition 1.3.2], see also [20]).

Proposition 2.13. *For any non-integer $\gamma > 0$, $C_B^\gamma = \mathcal{C}_B^\gamma$ and their norms are equivalent.*

Remark 2.14. We observe that, if $f \in C_B^\gamma$ with $\gamma \in (0,1]$, then there exist two positive constants C and $\nu \in (0,\gamma)$ (small enough depending on B) such that

$$\|f\|_{\mathcal{C}^\nu} := \|f\|_{L^\infty} + \sup_{z \neq z' \in \mathbb{R}^N,\ |z-z'| \leq 1} \frac{|f(z) - f(z')|}{|z - z'|^\nu} \leq C \|f\|_\gamma.$$

In other words, f is ν-Hölder continuous in the classical sense (see [3, Definition 1.49]).

The anisotropic Hölder spaces above can be somehow "extended" to functions defined on $\mathbb{R} \times \mathbb{R}^N$. These are the so-called *intrinsic* Hölder spaces mentioned in the introduction, which are the natural choices to include the regularity along the time-variable in the study of degenerate $\mathcal{K}$ and $\mathcal{K}'$-type operators. In this framework, the correct translation to consider is not the Euclidean one used in the parabolic case, but the one with respect to the integral curves of Y. Intuitively, this follows from the fact that Y plays the role played by the partial derivative ∂_t in the parabolic case. For the same reason, the Lie derivative along Y, i.e.

$$Yu(t,z) := \lim_{h \to 0} \frac{u(t+h, e^{hB} z) - u(t,z)}{h} \tag{2.17}$$

counts as a second-order derivative. Once more, we provide the definition of these spaces only for some ranges of γ.

Definition 2.15. For $T > 0$, we denote by $C_{B,T}^\gamma$ the space of functions $u : [0,T] \times \mathbb{R}^N \longrightarrow \mathbb{R}$ such that the following exists and it is finite (in particular it is a norm):

$$\|u\|_{C_{B,T}^\gamma} := \begin{cases} \sup_{(t,z) \in [0,T] \times \mathbb{R}^N} |u(t,z)| + \sup_{(t,z),(s,y) \in [0,T] \times \mathbb{R}^N} \frac{|u(t,z) - u(s,y)|}{|t-s|^{\gamma/2} + |z - e^{(t-s)B} y|_B^\gamma}, & \gamma \in (0,1], \\ \sup_{(t,z) \in [0,T] \times \mathbb{R}^N} |u(t,z)| + \sup_{\substack{(t,z) \in [0,T] \times \mathbb{R}^N \\ s \in [0,T]}} \frac{|u(s, e^{(s-t)B} z) - u(t,z)|}{|s-t|^{\gamma/2}} + \sum_{l=1}^d \|\partial_{z_l} u\|_{C_{B,T}^{\gamma-1}}, & \gamma \in (1,2], \\ \sup_{(t,z) \in [0,T] \times \mathbb{R}^N} |u(t,z)| + \sum_{l=1}^d \|\partial_{z_l} u\|_{C_{B,T}^{\gamma-1}} + \|Yu\|_{L_T^\infty \mathcal{C}_B^{\gamma-2}}, & \gamma \in (2,3]. \end{cases}$$

In the case $\gamma \in (2,3]$, we also require that Yu, which is the Lie derivative of u in the sense of (2.17), is continuous on $[0,T] \times \mathbb{R}^N$. Note that, by the boundedness of the norm, $u, \partial_{z_i} u, \partial_{z_i z_j} u$ are uniformly continuous on $[0,T] \times \mathbb{R}^N$. We say that these functions are instrinsically γ-Hölder continuous.

Obviously, by Proposition 2.13, for any $\gamma \in (0,1)$ we have

$$C_{B,T}^\gamma \subset C_T \mathcal{C}_b \cap L_T^\infty \mathcal{C}_B^\gamma. \tag{2.18}$$

A full characterization of the intrinsic Hölder spaces at any order was given in [52] (see also the references therein for previous characterizations), where an intrinsic Taylor formula was proved. The latter allows to see that (2.18) actually holds for any non-integer $\gamma > 0$.

2.4 Regularizations

Let Φ be a standard mollifier on $\mathbb{R}^N$ with compact support. For $n \in \mathbb{N}$ we set $\Phi_n(z) = n^N \Phi(nz)$. For $\gamma \in \mathbb{R}$, $g \in \mathcal{C}_B^\gamma$, set the family

$$g^{(n)} := \Phi_n * g, \qquad n \in \mathbb{N}. \tag{2.19}$$

Lemma 2.16. *Let $T > 0$, $g \in L_T^\infty \mathcal{C}_B^\gamma$ with $\gamma \in \mathbb{R}$ (respectively $g \in C_T \mathcal{C}_B^\gamma$), and denote by $g^{(n)}$ the function $([0,T] \ni t \mapsto g_t^{(n)})$ given by (2.19). Then we have the following.*

(i) For any $n \in \mathbb{N}$ and $\alpha > 0$, we have $g^{(n)} \in L_T^\infty \mathcal{C}_B^\alpha$ (respectively $g^{(n)} \in C_T \mathcal{C}_b \cap L_T^\infty \mathcal{C}_B^\alpha$).

(ii) For any $\eta \in (0,1)$, there exists $C = C(\gamma, \eta) > 0$ such that

$$\|g^{(n)} - g\|_{T,\gamma-\eta} \le C \|g\|_{T,\gamma} \int_{\mathbb{R}^N} \left|\frac{y}{n}\right|_B^\eta \Phi(y) dy, \qquad n \in \mathbb{N}, \tag{2.20}$$

and in particular,

$$\|g^{(n)} - g\|_{T,\gamma-\eta} \to 0, \qquad \text{as } n \to \infty.$$

(iii) In the case $g \in C_T \mathcal{C}_b$ with compact support, then $g^{(n)} \in C_T \mathcal{S}$ for any $n \in \mathbb{N}$.

To prove Lemma 2.16, we need the following result, which is a particular case of [33, Corollary 2.9].

Lemma 2.17. *Let $g \in \mathcal{C}_B^\gamma$ with $\gamma \in \mathbb{R}$, and $\eta \in (0,1)$. There exists a constant $C = C(\gamma, \eta) > 0$ such that*

$$\|\delta_h g\|_{\gamma-\eta} \le C |h|_B^\eta \|g\|_\gamma, \qquad h \in \mathbb{R}^N, \tag{2.21}$$

with $\delta_h g \in \mathcal{C}_B^\gamma$ being defined through

$$\langle \delta_h g | \phi \rangle := \langle g | \phi(\cdot - h) - \phi \rangle, \qquad \phi \in \mathcal{S}(\mathbb{R}^N).$$

Proof of Lemma 2.16. Let $n \in \mathbb{N}$ and $t \in [0,T]$ be fixed throughout the proof and dropped for ease of notation. To prove Part (i) it is sufficient to show that all $g^{(n)}$ are smooth in space with bounded derivatives uniformly in t. Smoothness is proved as in [35, Theorem 4.1.1]. We need to prove that for every multi-index $k \in \mathbb{N}^N$

$$|\partial^k g^{(n)}(z)| = |g * \partial^k \Phi_n(z)| = |\langle g | \, \partial^k \Phi_n(z - \cdot) \rangle| < \infty, \qquad z \in \mathbb{R}^N.$$

If $\varphi_z(y) := \partial^k \Phi_n(z - y)$, it follows from Remark 2.1-(c) that

$$\partial^k g^{(n)}(z) = \sum_{i,j \ge -1} \langle \Delta_i g | \, \Delta_j \varphi_z \rangle = \sum_{i \ge -1} \langle \Delta_i g | \, \widetilde{\Delta}_i \varphi_z \rangle$$

with

$$\widetilde{\Delta}_i \varphi_z := \begin{cases} (\Delta_{-1} + \Delta_0)\varphi_z, & i = -1, \\ (\Delta_{i-1} + \Delta_i + \Delta_{i+1})\varphi_z, & i \ge 0. \end{cases}$$

Therefore, using Hölder inequality we get

$$|\partial^k g^{(n)}(z)| \le \sum_{i \ge -1} \|\Delta_i g\|_{L^\infty} \|\widetilde{\Delta}_i \varphi_z\|_{L^1} = \sum_{i \ge -1} 2^{i\gamma} \|\Delta_i g\|_{L^\infty} 2^{-i\gamma} \|\widetilde{\Delta}_i \varphi_z\|_{L^1} \le \|g\|_\gamma \sum_{i \ge -1} 2^{-i\gamma} \underbrace{\|\widetilde{\Delta}_i \varphi_z\|_{L^1}}_{= \|\widetilde{\Delta}_i \varphi_0\|_{L^1}}. \tag{2.22}$$

Now, the latter sum is finite as it is bounded by the dual Besov norm $\sum_{i\geq -1} 2^{-i\gamma}\|\Delta_i\varphi_0\|_{L^1}$, which is finite due to $\varphi \in C_c^\infty$. Respectively, under the additional assumption $g \in C_T\mathcal{C}_B^\gamma$, with $k = 0$ in (2.22) and replacing $g^{(n)}$ with $g_t^{(n)} - g_s^{(n)}$ yields $g^{(n)} \in C_T\mathcal{C}_b$.

We now prove Part (ii). For any $j \geq -1$ we have

$$\Delta_j g_t^{(n)} = \check{\rho}_j * \Phi_n * g_t = \Phi_n * \check{\rho}_j * g_t = \Phi_n * \Delta_j g_t,$$

and thus

$$\begin{aligned}\big(\Delta_j\,(g_t^{(n)} - g_t)\big)(z) &= \int_{\mathbb{R}^N} (\Delta_j g_t)(z-y)\Phi_n(y)dy - \int_{\mathbb{R}^N} (\Delta_j g_t)(z)\Phi(y)dy \\ &= \int_{\mathbb{R}^N} \underbrace{\big((\Delta_j g_t)(z - y/n) - (\Delta_j g_t)(z)\big)}_{=\delta_{y/n}(\Delta_j g_t)(z)} \Phi(y)dy, \qquad z \in \mathbb{R}^N,\end{aligned}$$

which yields

$$\|\Delta_j\,(g_t^{(n)} - g_t)\|_{L^\infty} \leq \int_{\mathbb{R}^N} \|\delta_{y/n}(\Delta_j g_t)\|_{L^\infty}\Phi(y)dy$$

(by (2.21) and definition of Besov norm)

$$\leq C\|g\|_{T,\gamma}\, 2^{-j(\gamma-\eta)} \int_{\mathbb{R}^N} \Big|\frac{y}{n}\Big|_B^\eta \Phi(y)dy.$$

Again by definition of Besov norm, the latter inequality implies (2.20) and concludes the proof.

We now prove Part (iii). It is sufficient to prove that $t \mapsto \mathcal{F}g^{(n)}(t,\cdot)$ is continuous with values in $\mathcal{S}$. We have

$$\mathcal{F}g^{(n)}(t,\cdot) = \mathcal{F}g(t,\cdot)\,\mathcal{F}\Phi_n.$$

Since g is a continuous function with compact support then $(t \mapsto \mathcal{F}g(t,\cdot)) \in C_T\mathcal{S}$. Since $\mathcal{F}\Phi_n \in \mathcal{S}$ the result easily follows. □

The next result will be used in Section 5.1 to prove the equivalence between our formulation of the martingale problem with the classical one by Stroock-Varadhan.

Lemma 2.18. *Let $T > 0$, $\alpha \in (0,1]$. For any $u \in C_{B,T}^{2+\alpha}$, there exists a sequence $(u^{(n)})_{n\in\mathbb{N}}$ such that the following hold.*

(a) For any $n \in \mathbb{N}$, $u^{(n)} \in C^{1,2}([0,T]\times\mathbb{R}^N)$, with $u^{(n)}, \partial_{z_i}u^{(n)}, \partial_{z_iz_j}u^{(n)}$ bounded on $[0,T]\times\mathbb{R}^N$.

(b) For any $i,j = 1,\cdots,d$ and for any compact $K \subset \mathbb{R}^N$, we have

$$\|u - u^{(n)}\|_T,\ \|\partial_{z_i}u - \partial_{z_i}u^{(n)}\|_T,\ \|\partial_{z_iz_j}u - \partial_{z_iz_j}u^{(n)}\|_T,\ \|(Yu - Yu^{(n)})\mathbf{1}_K\|_T \longrightarrow 0, \qquad \text{as } n \to +\infty, \tag{2.23}$$

where Yu denotes the Lie derivative of u, continuous on $[0,T]\times\mathbb{R}^N$, defined as in (2.17).

Proof. Let Φ be a standard mollifier on $\mathbb{R}^N$ and set, for any $n \in \mathbb{N}$,

$$u^{(n)}(t,z) := \frac{1}{|\det e^{tB}|}\int_{\mathbb{R}^N} \Phi_n\big(e^{-tB}(z-y)\big)u(t,y)dy, \quad \Phi_n(y) := n^N\Phi(ny), \qquad (t,z) \in [0,T]\times\mathbb{R}^N.$$

We first prove Part (a). By the properties of the mollifier and by the fact that u is continuous and bounded on $[0,T]\times\mathbb{R}^N$, it is standard to prove that $\nabla_z u^{(n)}(t,z)$ is continuous and bounded on $[0,T]\times\mathbb{R}^N$

and that the same holds true for the second order derivatives in space. We now study $\partial_t u^{(n)}(t,z)$. First note that a simple change of variable yields

$$u^{(n)}(t,z) = \int_{\mathbb{R}^N} \Phi_n(y) u(t, z - e^{tB} y) dy. \tag{2.24}$$

Note that the Lie derivative along Y is invariant with respect to the translation along its integral curves, namely

$$\begin{aligned} Y\big((t,z) \mapsto u\big(t, z - e^{tB}y\big)\big) &= \lim_{\delta \to 0} \frac{u\big(t+\delta, e^{\delta B}(z - e^{tB}y)\big) - u\big(t, z - e^{tB}y\big)}{\delta} \\ &= \lim_{\delta \to 0} \frac{u\big(t+\delta, e^{\delta B}\xi\big) - u(t,\xi)}{\delta}\Big|_{\xi = z - e^{tB}y} = (Yu)\big(t, z - e^{tB}y\big), \end{aligned}$$

for any $0 \le t \le T$ and $z, y \in \mathbb{R}^N$. Owing to this, and to the boundedness of Yu on $[0,T] \times \mathbb{R}^N$, one can move the Lie derivative under the integral in (2.24) and obtain

$$Yu^{(n)}(t,z) = \int_{\mathbb{R}^N} \Phi_n(y) (Yu)\big(t, z - e^{tB}y\big) dy. \tag{2.25}$$

Therefore, $u^{(n)}(t,z)$ is differentiable with respect to t and we have

$$\partial_t u^{(n)}(t,z) = Yu^{(n)}(t,z) - \langle Bz, \nabla_z u^{(n)}(t,z) \rangle, \qquad (t,z) \in [0,T] \times \mathbb{R}^N,$$

where $\nabla_z u^{(n)}(t,z)$ makes sense a standard Euclidean gradient. As $Yu^{(n)}, \nabla_z u^{(n)}$ are continuous on $[0,T] \times \mathbb{R}^N$, so is $\partial_t u^{(n)}$. This concludes the proof of Part (a).

We now prove Part (b). By the boundedness of $\partial_{z_i} u, \partial_{z_i z_j} u$ on $[0,T] \times \mathbb{R}^N$, one can move the derivatives under the sign of the integral in (2.24) and obtain

$$\partial_{z_i} u^{(n)}(t,z) = \int_{\mathbb{R}^N} \Phi_n(y) \partial_{z_i} u\big(t, z - e^{tB}y\big) dy, \tag{2.26}$$

$$\partial_{z_i z_j} u^{(n)}(t,z) = \int_{\mathbb{R}^N} \Phi_n(y) \partial_{z_i z_j} u\big(t, z - e^{tB}y\big) dy. \tag{2.27}$$

Noticing now that, by assumption, $u, \partial_{z_i} u, \partial_{z_i z_j} u$ are uniformly continuous on $[0,T] \times \mathbb{R}^N$ and Yu is uniformly continuous on $[0,T] \times K$ for any compact $K \subset \mathbb{R}^N$, it is standard to show that (2.24), (2.25), (2.26) and (2.27) imply (2.23). □

3 Non-linear Fokker-Planck singular PDE

In this section we study the Fokker-Planck PDE associated to the McKean-Vlasov SDE (1.3), i.e.

$$\begin{cases} \mathcal{K}' u_t = \mathrm{div}_v\big(u_t F(u_t) b_t\big), & t \in (0,T], \\ u_0 = u^0, \end{cases} \tag{3.1}$$

with $\mathcal{K}'$ being the forward Kolmogorov operator associated to the Markovian SDE (2.11), defined in (2.13)-(2.14), and where div_v is the divergence operator in $\mathbb{R}^d$. Namely div_v formally acts on a function $G : \mathbb{R}^N \to \mathbb{R}^d$ as

$$\mathrm{div}_v G(z) = \sum_{j=1}^{d} \partial_{v_j} G_j(v,x), \qquad z = (v,x) \in \mathbb{R}^d \times \mathbb{R}^{N-d}.$$

We assume, throughout this section, that Assumptions 1.1, 1.4 and 1.5 for $\Phi = F$ and $\Phi = \widetilde{F}$ (see (1.13)) are satisfied.

3.1 Well-posedeness of the PDE

We prove that (3.1) has a unique mild (or, equivalently, weak) solution in $C_T\mathcal{C}_b \cap L_T^\infty \mathcal{C}_B^{\beta+\varepsilon}$ for a suitable $\varepsilon > 0$. Note that, a priori, it is not obvious how the term

$$\mathrm{div}_v(u_t F(u_t) b_t), \qquad t \in [0,T],$$

in (3.1) might be well-defined. In this respect and to prove well-posedness of (3.1) we need the following

Lemma 3.1. *Let Φ satisfy Assumption 1.5. For any $\alpha \in (0,1)$, the map*

$$\Phi : \mathcal{C}_B^\alpha \longrightarrow \mathcal{C}_B^\alpha, \qquad f \longmapsto \Phi f := \Phi \circ f,$$

is well-defined and there exists a constant $C > 0$, only depending on Φ, such that

$$\|\Phi f - \Phi g\|_\alpha \le C(1 + \|f\|_\alpha + \|g\|_\alpha)\|f - g\|_\alpha, \qquad f, g \in \mathcal{C}_B^\alpha, \tag{3.2}$$

$$\|\Phi f\|_\alpha \le C(1 + \|f\|_\alpha), \qquad f \in \mathcal{C}_B^\alpha. \tag{3.3}$$

Proof. We only show (3.2), with (3.3) being simpler. Furthermore, without loosing generality, we can assume Φ being scalar valued. As the function Φ is Lipschitz continuous, we have

$$\|\Phi f - \Phi g\|_{L^\infty} \le C\|f - g\|_{L^\infty}.$$

Now fix $z, z' \in \mathbb{R}^d$. To complete the proof of (3.2), we need to show

$$|\Phi f(z) - \Phi g(z) - \Phi f(z') + \Phi g(z')| \le C(1 + \|f\|_\alpha + \|g\|_\alpha)\|f - g\|_\alpha |z - z'|_B^\alpha. \tag{3.4}$$

We set

$$a := f(z) - g(z), \quad b := f(z') - g(z'), \quad c := f(z) - f(z'), \quad d := g(z) - g(z'),$$

and proceed by considering two separate cases:

$\boxed{|a| + |b| \le |c| + |d|}$ By mean-value theorem, there exists ξ included between $f(z)$ and $g(z)$, and η included between $f(z')$ and $g(z')$, such that

$$\Phi f(z) - \Phi g(z) = \Phi'(\xi) a, \qquad \Phi f(z') - \Phi g(z') = \Phi'(\eta) b.$$

Therefore, by triangular inequality we obtain

$$|\Phi f(z) - \Phi g(z) - \Phi f(z') + \Phi g(z')| \le |\Phi'(\xi)(a - b)| + \big|b\big(\Phi'(\xi) - \Phi'(\eta)\big)\big|$$

(by Assumption 1.5)

$$\le C\|f - g\|_\alpha \big(|z - z'|_B^\alpha + |\xi - \eta|\big). \tag{3.5}$$

Now, as $|a| + |b| \le |c| + |d|$, we obtain

$$|\xi - \eta| \le |a| + |b| + |c| + |d| \le 2(|c| + |d|) \le 2(\|f\|_\alpha + \|g\|_\alpha)|z - z'|_B^\alpha,$$

which, together with (3.5), yields (3.4).

$\boxed{|c|+|d| \le |a|+|b|}$ By mean-value theorem, there exist ξ included between $f(z)$ and $f(z')$, and η included between $g(z)$ and $g(z')$, such that

$$\Phi f(z) - \Phi f(z') = \Phi'(\xi)c, \qquad \Phi g(z) - \Phi g(z') = \Phi'(\eta)d.$$

Therefore, by triangular inequality we obtain

$$|\Phi f(z) - \Phi g(z) - \Phi f(z') + \Phi g(z')| \le |\Phi'(\xi)(c-d)| + \big|d\big(\Phi'(\xi) - \Phi'(\eta)\big)\big|$$

(by Assumption 1.5)

$$\le C\big(\|f-g\|_\alpha + \|g\|_\alpha|\xi-\eta|\big)|z-z'|_B^\alpha. \tag{3.6}$$

Now, as $|c|+|d| \le |a|+|b|$, we obtain

$$|\xi-\eta| \le |a|+|b|+|c|+|d| \le 2(|a|+|b|) \le 4\|f-g\|_\alpha,$$

which, together with (3.6), yields (3.4) and completes the proof. □

Remark 3.2. By Lemma 3.1 applied to $\Phi = \widetilde{F}$, for any $\alpha \in (\beta, 1)$ we have

$$\mathrm{div}_v(fF(f)b_t) \in \mathcal{C}_B^{-\beta-1}, \qquad f \in C_B^\alpha, \quad t \in [0,T].$$

Indeed, by the Bernstein inequality (2.8), the estimate for the Bony's product (2.10) and by (3.3), we obtain

$$\|\mathrm{div}_v(fF(f)b_t)\|_{-\beta-1} \le C\sum_{i=1}^d \big\|\big((\widetilde{F}f)b_t\big)_i\big\|_{-\beta} \le C(1+\|f\|_\alpha)\|b_t\|_{-\beta} < +\infty. \tag{3.7}$$

Definition 3.3. We say that a function $u \in C_T\mathcal{C}_b \cap L_T^\infty \mathcal{C}_B^{\beta+\varepsilon}$ is a *weak solution* for the singular Fokker-Planck (3.1) if it is a distributional solution. Namely, for all $\varphi \in \mathcal{S}$ we have

$$\langle u_t|\,\varphi\rangle = \langle u^0|\,\varphi\rangle + \int_0^t \langle u_s|\,\mathcal{A}\,\varphi\rangle\,ds + \sum_{i=1}^d \int_0^t \big\langle u_s(F(u_s)b_s)_i|\,\partial_{v_i}\varphi\big\rangle ds, \qquad t \in [0,T],$$

where $\mathcal{A}$ is the generator of Z in (2.11), defined through (1.9).

We now introduce the notion of *mild solution*, which turns out to be equivalent to the one of *weak solution*.

Definition 3.4. We say that a function $u \in C_T\mathcal{C}_b \cap L_T^\infty \mathcal{C}_B^{\beta+\varepsilon}$ is a *mild solution* for the singular Fokker-Planck equation (3.1) if the integral equation

$$u_t = P'_t u^0 - \int_0^t P'_{t-s}\big(\mathrm{div}_v\,(u_s F(u_s)\,b_s)\,\big)ds, \qquad t \in [0,T], \tag{3.8}$$

is satisfied.

Remark 3.5. Inequality (3.7) in Remark 3.2 together with the Schauder's estimate (2.16) yield

$$\big\|P'_{t-s}\big(\mathrm{div}_v(u_sF(u_s)\,b_s)\big)\big\|_{\beta+\varepsilon} \le C(t-s)^{-(\beta+\frac{\varepsilon+1}{2})}\|\mathrm{div}_v(u_sF(u_s)\,b_s)\|_{-\beta-1} \le C(t-s)^{-(\beta+\frac{\varepsilon+1}{2})}, \tag{3.9}$$

where

$$\beta + \frac{\varepsilon+1}{2} < 1,$$

in light of $\beta \in (0,1/2)$ and $\varepsilon \in (0, 1-2\beta)$. This shows that the LHS in (3.9) is integrable as a function of s, in $[0,t]$, and thus Definition 3.4 is well-posed. Indeed, the integral in (3.8) can be understood as a pointwise Lebesgue integral, namely as

$$\int_0^t P'_{t-s}\big(\mathrm{div}_v\,(u_s F(u_s)\, b_s)\,\big)(x)ds, \qquad x \in \mathbb{R}^N,$$

and the standard triangular inequality yields

$$\|u_t\|_{\beta+\varepsilon} \le \|P'_t u^0\|_{\beta+\varepsilon} + \int_0^t \big\|P'_{t-s}\big(\mathrm{div}_v(u_s F(u_s)\, b_s)\big)\big\|_{\beta+\varepsilon} ds.$$

Remark 3.6.

1. The right-hand side of (3.8) can also be considered as a Bochner's integral with values in $\mathcal{S}'$ equipped with its strong topology, being the strong dual of the vector space $\mathcal{S}$, see e.g. [62, Chapter 4, Example 1 a]. For $k \in \mathbb{N}$, let $(E_k, \|\cdot\|_k)$ be the normed space

$$E_k := \{\varphi \in C^\infty(\mathbb{R}^N) |\, \|\varphi\|_k := \sup_{z\in\mathbb{R}^N,\, |\alpha|\le k} |D^\alpha \varphi(z)|\, \big(|z|^2+1\big)^k < \infty\}. \tag{3.10}$$

   Then $\mathcal{S}$ is a vector (metrizable) space given by the intersection of the E_k's, while $\mathcal{S}'$ coincides with the union of the spaces E_k^*, which are the strong duals of $(E_k, \|\cdot\|_k)$. The (strong) topology of $\mathcal{S}'$ is the one corresponding with the inductive limit of the E_k^*'s.

2. Since E_k^* endowed with the strong topology is a Banach space, it is then possible to define the Bochner's integral $\int_0^T g_s ds$ for $g \in L_T^\infty \mathcal{S}'$. The latter is the space of strongly measurable $g : [0,T] \to \mathcal{S}'$ such that there is k with $g(t) \in E_k^*$ for a.e. $t \in [0,T]$ and that $t \mapsto \|g(t)\|_{E_k^*}$ belongs to $L^\infty([0,T])$.

3. Specifically, if $g \in L_T^\infty \mathcal{S}'$ then we have the following.

   i) The map $t \mapsto \langle u_t |\, \varphi\rangle$ is Borel for all $\varphi \in \mathcal{S}$.

   ii) There exists $C = C(g) > 0$, $k \in \mathbb{N}$ such that for all $\varphi \in \mathcal{S}$

$$|\langle g_t |\, \varphi\rangle| \le C\|\varphi\|_k \qquad \text{a.e. in } [0,T].$$

4. We say that $g : [0,T] \to \mathcal{S}'$ belongs to $C^{AC}([0,T];\mathcal{S}')$ if there is $g' : [0,T] \to \mathcal{S}'$ continuous with $g(t) = g(0) + \int_0^t g'(s)ds$ for any $t \in [0,T]$.

The next result shows the equivalence between mild and weak solutions. Although the statement assumes $u \in C_T\mathcal{C}_b \cap L_T^\infty \mathcal{C}_B^{\beta+\varepsilon}$ in accordance with Definitions 3.3 and 3.4, its proof relies on the weaker assumption that $u \in L_T^\infty \mathcal{S}'$, the latter being the natural framework in order to make sense of the the concepts of mild and weak solutions to the linear version of the Fokker-Planck equation.

Proposition 3.7. *Let $u \in C_T\mathcal{C}_b \cap L_T^\infty \mathcal{C}_B^{\beta+\varepsilon}$. Then u is a weak solution of* (3.1) *if and only if it is a mild solution.*

Proof. For $g \in L_T^\infty \mathcal{S}'$, $t \in [0,T]$, and $\varphi \in \mathcal{S}$ we have

$$|\langle P'_{t-s} g_s |\, \varphi\rangle| = |\langle g_s |\, P_{t-s}\varphi\rangle| \le C(g)\|\varphi\|_k, \qquad s \in [0,t],$$

for some $k \in \mathbb{N}$, where $\|\cdot\|_k$ was defined in (3.10). Then $s \mapsto P'_{t-s} g_s \in L_t^\infty \mathcal{S}'$ and, for $f_0 \in \mathcal{S}'$, the mild function $u : [0,T] \to \mathcal{S}'$ defined by

$$u_t := P'_t u^0 - \int_0^t P'_{t-s} g_s ds, \qquad t \in [0,T],$$

is well-defined in $L^\infty_T\mathcal{S}'$. By inspection, it can be verified that such a u is a weak solution of

$$\begin{cases} \mathcal{K}'\, u_t = g_t, & t \in (0,T), \\ u_0 = u^0, \end{cases} \tag{3.11}$$

namely, for every $\varphi \in \mathcal{S}$

$$\langle u_t |\, \varphi\rangle = \langle u^0 |\, \varphi\rangle + \int_0^t \big(\langle u_s |\, \mathcal{A}\, \varphi\rangle + \langle g_s |\, \varphi\rangle\big) ds, \qquad t \in [0,T].$$

By setting $g_s := \mathrm{div}_v(u_s F(u_s) b_s)$, it holds that a mild solution of (3.1) is also a weak solution. To prove the opposite implication it is then sufficient to prove uniqueness of weak solutions to (3.11) in the sense above. This fact can be established in the same way as in [41, Proposition 3.18], even though that result was stated for finite measure-valued functions. □

We now state one of the main results of this section.

Theorem 3.8. *Let Assumptions 1.1, 1.4, and Assumption 1.5 for $\Phi = \widetilde{F}$, hold. For all $\varepsilon \in (0, 1-2\beta)$ such that $u^0 \in \mathcal{C}^{\beta+\varepsilon}_B$ there exists a unique mild solution (or equivalently, by Proposition 3.7, a weak solution) $u \in C_T\mathcal{C}_b \cap L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$ to* (3.1).

Remark 3.9. Notice that, given an initial condition $u^0 \in \mathcal{C}^{\beta+\varepsilon}_B$ for some $\varepsilon \in (0, 1-2\beta)$, the solution u to equation (3.1) exists in the space $C_T\mathcal{C}_b \cap L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$ but it is unique in any larger space of the form $C_T\mathcal{C}_b \cap L^\infty_T\mathcal{C}^{\beta+\varepsilon'}_B$ for $0 < \varepsilon' < \varepsilon$.

Remark 3.10. The result of Theorem 3.8 remains true, with identical proof, under a weaker version of Assumption 1.4, namely replacing $b \in C_T\mathcal{C}^{-\beta}_B$ by $b \in L^\infty_T\mathcal{C}^{-\beta}_B$.

To prove Theorem 3.8, the idea is to find a fixed point $w^* \in L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$ for the map

$$J_t(w) := -\int_0^t P'_{t-s}[\mathrm{div}_v G_s(w)] ds, \qquad G_s(w) := \widetilde{F}(w_s + P'_s u^0) b_s, \qquad t \in [0,T]. \tag{3.12}$$

A mild solution to (3.1) can be then obtained by setting $u_t := w^*_t + P'_t u^0$, after verifying that $u \in C_T\mathcal{C}_b$. The uniqueness of the solution will stem from the fact that the fixed point w^* is proved via contraction.

Remark 3.11. Proceeding as in Remark 3.2 one can see that, if $w \in L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$, the integral in (3.12) can be understood as a standard pointwise Lebesgue integral, namely

$$J_t(w)(x) = -\int_0^t P'_{t-s}[\mathrm{div}_v G_s(w)](x)\, ds, \qquad x \in \mathbb{R}^N.$$

The standard triangular inequality also yields

$$\|J_t(w)\|_{\beta+\varepsilon} \le \int_0^t \|P'_{t-s}[\mathrm{div}_v G_s(w)]\|_{\beta+\varepsilon}\, ds.$$

For any $\rho \ge 0$ we define the norm on $L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$ given by

$$\|w\|_{\rho,T,\beta+\varepsilon} := \sup_{t\in[0,T]} e^{-\rho t}\|w_t\|_{\beta+\varepsilon}, \tag{3.13}$$

and we consider the ρ-closed balls

$$E^{\beta+\varepsilon}_{\rho,M} := \{w \in L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B \,:\, \|w\|_{\rho,T,\beta+\varepsilon} \le M\}, \qquad M > 0.$$

Remark 3.12. For any $\rho \geq 0$, the norm $\|\cdot\|_{\rho,T,\beta+\varepsilon}$ is equivalent to $\|\cdot\|_{T,\beta+\varepsilon}$. Furthermore, the ρ-closed balls are closed sets of $L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$, and thus $(E^{\beta+\varepsilon}_{\rho,M}, \|\cdot\|_{\rho,T,\beta+\varepsilon})$ is a complete metric space for any $M > 0$.

The choice of these spaces is justified by the following lemma.

Lemma 3.13. *There exist $\rho_0, M_0 > 0$, depending on T, β, ε, $\|u^0\|_{\beta+\varepsilon}$, $\|b\|_{T,-\beta}$, such that*

$$J : E^{\beta+\varepsilon}_{\rho_0,M} \longrightarrow E^{\beta+\varepsilon}_{\rho_0,M}, \qquad M > M_0.$$

Furthermore, for any $M > M_0$ there exists $\rho_M > 0$, also dependent on T, β, ε, B, $\|u^0\|_{\beta+\varepsilon}$, $\|b\|_{T,-\beta}$, such that

$$\|J(w) - J(w')\|_{\rho_M,T,\beta+\varepsilon} \leq \frac{1}{2}\|w - w'\|_{\rho_M,T,\beta+\varepsilon}, \qquad w, w' \in E^{\beta+\varepsilon}_{\rho_0,M}.$$

Proof. Owing to the anisotropic Schauder's estimate of Theorem 2.10, the proof is identical to the one of [40, Proposition 3.5 and Lemma 3.6]. □

Lemma 3.14. *We have*

$$(t \mapsto P'_t g) \in C_T\mathcal{C}_b, \qquad g \in \mathcal{C}_b, \tag{3.14}$$

$$J(w) \in C_T\mathcal{C}_b, \qquad w \in L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B. \tag{3.15}$$

Proof. The proof of (3.14) is standard: indeed $\Gamma(t-s,y,z)$ is smooth on $\{s > t\}$, and (2.15) holds true uniformly with respect to $z \in \mathbb{R}^N$. Let now $w \in L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$ be fixed. Lemma 3.13 implies that $J_t(w) \in \mathcal{C}_b$ for any $t \in [0,T]$. To show (3.15), note that applying (3.7) in Remark 3.2 with $f = w_s + P'_s u^0 \in \mathcal{C}^{\beta+\varepsilon}_B$ and Schauder's estimate (2.16) with $\alpha = 0$ yields

$$\|\mathrm{div}_v\, G_s(w)\|_{-\beta-1} \leq C(1 + \|w_s\|_{\beta+\varepsilon} + \|u^0\|_{\beta+\varepsilon})\|b_s\|_{-\beta}, \qquad s \in [0,T],$$

and thus $\mathrm{div}_v G(w) \in L^\infty_T\mathcal{C}^{-\beta-1}_B$. Therefore, it is sufficient to show that

$$t \longmapsto \int_0^t P'_{t-s} f_s ds \in C_T\mathcal{C}_b, \qquad f \in L^\infty_T\mathcal{C}^{-\beta-1}_B.$$

We study the continuity in time, by splitting the integrals as follows. For $h \in (0, T-t)$, using the semigroup property, we have

$$\begin{aligned}\int_0^{t+h} P'_{t+h-s} f_s ds - \int_0^t P'_{t-s} f_s ds &= \int_0^t \Big(P'_h P'_{t-s} f_s - \underbrace{P'_{t-s} f_s}_{=:g_s}\Big) + \int_t^{t+h} P'_{t+h-s} f_s ds \\ &= \int_0^t (P'_h g_s - g_s)\, ds + \int_t^{t+h} P'_{t+h-s} f_s ds =: I^h_1 + I^h_2.\end{aligned}$$

To estimate $\|I^h_2\|_\infty$, we set $\delta \in (0, \frac{1}{2})$. By the Schauder's estimate (2.16) we obtain

$$\|I^h_2\|_\infty \leq \int_t^{t+h} \|P'_{t+h-s} f_s\|_\infty ds \leq C \int_t^{t+h} \|P'_{t+h-s} f_s\|_\delta\, ds \leq C\|f\|_{T,-\beta-1} \int_0^h (h-s)^{-\frac{\beta+1+\delta}{2}} ds \xrightarrow[h\to 0^+]{} 0.$$

To estimate $\|I^h_1\|_\infty$, note that (3.14) yields

$$\|P'_h g_s - g_s\|_\infty \to 0, \qquad \text{as} \quad h \to 0^+,$$

for any $s \in [0,t)$. Also, $\|P'_h g_s\|_\infty \leq C\|g_s\|_\infty$, the latter being summable on $[0,t]$ (once more by Schauder's estimate (2.16)). Therefore, Lebesgue's dominated convergence theorem yields $\|I^h_1\|_\infty \to 0$ as $h \to 0^+$. This proves continuity from the right. Continuity from the left can be shown with analogous computations. □

We are now in the position to prove Theorem 3.8.

Proof of Theorem 3.8. Since $P.u^0 \in L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$, a function $u \in L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$ satisfies (3.8) if and only if $u_t = w^*_t + P_t u^0$ with

$$J(w^*) = w^* \in L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B.$$

Lemma 3.13 and Remark 3.12 allow us to apply Banach fixed point theorem and obtain that such w^* exists and is unique. Finally, Lemma 3.14 ensures that both $(t \mapsto P_t u^0), w^* \in C_T\mathcal{C}_b$, and this completes the proof. □

3.2 The regularized equation

For any $n \in \mathbb{N}$, consider the regularized Fokker-Planck Cauchy problem

$$\begin{cases} \mathcal{K}' u_t = \mathrm{div}_v\big(u_t F(u_t) b^{(n)}_t\big), & t \in (0,T], \\ u_0 = u^0, \end{cases} \tag{3.16}$$

where $b^{(n)}$ is the smooth approximation of b defined in (2.19).

Theorem 3.15. *Let Assumptions 1.1, 1.4 with related ε, and 1.5 for $\Phi = \widetilde{F}$ be in force. Then the following properties hold.*

(a) For any $n \in \mathbb{N}$ and $\alpha > 0$, we have $b^{(n)} \in C_T\mathcal{C}_b \cap L^\infty_T\mathcal{C}^\alpha_B$, and the regularized Fokker-Planck Cauchy problem (3.16) *admits a unique mild solution $u^{(n)} \in C_T\mathcal{C}_b \cap L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$, in the sense of Definition 3.4, with*

$$u^{(n)}_t \geq 0, \qquad \|u^{(n)}_t\|_{L^1(\mathbb{R}^N)} \leq 1, \qquad t \in [0,T]. \tag{3.17}$$

(b) We have

$$\|u^{(n)} - u\|_{T,\beta+\varepsilon} \to 0, \qquad \text{as } n \to \infty, \tag{3.18}$$

where $u \in C_T\mathcal{C}_b \cap L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$ is the unique solution to (3.1).

The rest of the section is devoted to the proof of Theorem 3.15. We start by proving the first part.

Proof of Theorem 3.15, Part (a). Here n is fixed. By Lemma 2.16-(i), $b^{(n)} \in L^\infty_T\mathcal{C}^\alpha_B$ for any $\alpha > 0$. In particular, we have $b^{(n)} \in L^\infty_T\mathcal{C}^{-\beta}_B$, and, by Remark 3.10, the regularized FP Cauchy problem (3.16) admits a unique mild solution $u^{(n)}$ such that

$$u^{(n)} \in C_T\mathcal{C}_b \cap L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B. \tag{3.19}$$

We now show (3.17). Setting

$$\mathcal{B}^{(n)}_t(z) := F\big(u^{(n)}_t(z)\big) b^{(n)}_t(z), \qquad t \in [0,T], \quad z \in \mathbb{R}^N,$$

the function $u^{(n)}$ solves the linear FP Cauchy problem

$$\begin{cases} \mathcal{K}' u_t = \mathrm{div}_v\big(u_t \mathcal{B}^{(n)}_t\big), & t \in (0,T], \\ u_0 = u^0. \end{cases} \tag{3.20}$$

We can show that $\mathcal{B}^{(n)} \in C_T\mathcal{C}_b \cap L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$: by Lemma 2.16-(i), $b^{(n)} \in C_T\mathcal{C}_b \cap L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$, and $F\big(u^{(n)}_t(z)\big) \in L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$ thanks to Lemma 3.1 with $\Phi = F$, and $F\big(u^{(n)}_t(z)\big) \in C_T\mathcal{C}_b \cap L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$ thanks to (3.19), thus

$\mathcal{B}^{(n)} \in C_T\mathcal{C}_b \cap L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$. Since the PDE in (3.20) is written in divergence form, the results about classical solutions for degenerate operators in non-divergence form are not directly applicable since $\mathcal{B}^{(n)}$ is not regular enough. We thus need a further regularization step. For any $m \in \mathbb{N}$, consider the linear Cauchy problem

$$\begin{cases} \mathcal{K}' u_t - \mathrm{div}_v\big(u_t \mathcal{B}^{(n,m)}_t\big) = 0, \qquad t \in (0,T], \\ u_0 = u^0, \end{cases} \tag{3.21}$$

where $\mathcal{B}^{(n,m)}$ denotes the regular approximation of $\mathcal{B}^{(n,m)}$ defined as in Lemma 2.16. Obviously $\mathcal{B}^{(n,m)} \in C_T\mathcal{C}_b$ and by Lemma 2.16-(i), $\mathcal{B}^{(n,m)} \in C_T\mathcal{C}_b \cap L^\infty_T\mathcal{C}^\alpha_B$ for any $\alpha > 0$. Therefore, by [47, Theorem 2.7 and Remark 2.9], (3.21) has a strong Lie solution $u^{(n,m)}$, which can be represented (see [47, equation (3.1)]) as

$$u^{(n,m)}_t = \int_{\mathbb{R}^N} \Gamma^{(m)}(0,y;t,\cdot) u^0(y) dy, \qquad t \in [0,T], \tag{3.22}$$

where $\Gamma^{(m)}(s,y;t,z)$ is the fundamental solution of the operator

$$\mathcal{K}' - \langle \mathcal{B}^{(n,m)}, \nabla_v \rangle - \mathrm{div}_v \mathcal{B}^{(n,m)}$$

in the variables (t,z). Also, proceeding like in the proof of [24, Theorem 1.5], it can be shown that $\Gamma^{(m)}(s,y;t,z)$ is also the fundamental solution of the operator

$$\mathcal{K} + \langle \mathcal{B}^{(n,m)}, \nabla_v \rangle$$

in the variables (s,y). In particular $\Gamma^{(m)} > 0$ (see [46, Theorem 1.1]) and $\int_{\mathbb{R}^N} \Gamma^{(m)}(0,y;t,z)dz = 1$ (by [47, Theorem 2.7]). The latter and (3.22) imply (3.17) for $u^{(n,m)}$. Therefore, owing to the positivity of the functions and to Fatou's lemma, to show (3.17) for $u^{(n)}$ it is enough to prove

$$\|u^{(n,m)} - u^{(n)}\|_{T,\beta+\varepsilon} \to 0, \qquad \text{as } m \to \infty. \tag{3.23}$$

Owing once more to [47, (3.1)], it is straightforward to see that $u^{(n,m)}$ is also a mild solution to (3.21) in the sense of Definition 3.4, and the proof of (3.23) is a simpler version of the proof of (3.18): we omit the details for brevity. □

To prove the second part of Theorem 3.15, we need the following stability result.

Lemma 3.16. *Let the assumptions of Theorem 3.15 be in force. Let $\widetilde{b} \in L^\infty_T\mathcal{C}^{-\beta}_B$, and denote by $\widetilde{u} \in C_T\mathcal{C}_b \cap L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$ the mild solution of the nonlinear FP Cauchy problem obtained by replacing b with $\widetilde{b}$ in* (3.1)*, which exists in light of Remark 3.10. Then, for any $0 \le \eta < \min(\varepsilon, 1-2\beta-\varepsilon)$ there exists a function $l_{\varepsilon,\eta} : \mathbb{R}^+ \times \mathbb{R}^+ \longrightarrow \mathbb{R}^+$, independent of b and $\widetilde{b}$, increasing in the second variable, such that*

$$\|u - \widetilde{u}\|_{T,\beta+\varepsilon} \le l_{\varepsilon,\eta}\big(\|u^0\|_{\beta+\varepsilon}, \|b\|_{T,-\beta-\eta} \vee \|\widetilde{b}\|_{T,-\beta-\eta}\big)\|b - \widetilde{b}\|_{T,-\beta-\eta},$$

where $u \in C_T\mathcal{C}_b \cap L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$ is the unique solution to (3.1)*.*

Proof. Owing to the anisotropic Schauder's estimate of Lemma 2.10, and the continuity of the function $\widetilde{F}$ by Lemma 3.1, the proof is a straightforward modification of the one of [40, Proposition 4.7-(i)], upon substituting $\|u\|_{T,\alpha}$, $\|\widetilde{u}\|_{T,\alpha}$ with $\|u\|_{T,\beta+\varepsilon}$, $\|\widetilde{u}\|_{T,\beta+\varepsilon}$ and $\|b\|_{T,-\beta}$, $\|\widetilde{b}\|_{T,-\beta}$ with $\|b\|_{T,-\beta-\eta}$, $\|\widetilde{b}\|_{T,-\beta-\eta}$. □

We are now in the position to prove Part (b) of Theorem 3.15.

Proof of Theorem 3.15, Part (b). The result stems from combining Lemma 2.16-(ii) with Lemma 3.16 with $\widetilde{u} = u^{(n)}$ and $\widetilde{b} = b^{(n)}$ and using the fact that

$$l_{\varepsilon,\eta}(\|u^0\|_{\beta+\varepsilon}, \|b\|_{T,-\beta-\eta} \vee \|b^{(n)}\|_{T,-\beta-\eta}) \leq l_{\varepsilon,\eta}(\|u^0\|_{\beta+\varepsilon}, \|b\|_{T,-\beta-\eta} \vee \sup_n \|b^{(n)}\|_{T,-\beta-\eta}).$$

□

Remark 3.17. The result of Theorem 3.15 remains true, with identical proof, under a weaker version of Assumption 1.4, namely replacing $b \in C_T\mathcal{C}_B^{-\beta}$ by $b \in L_T^\infty\mathcal{C}_B^{-\beta}$.

4 Backward Kolmogorov PDE with singular drift

In this section we study the backward Kolmogorov Cauchy problem associated to the Markovian kinetic SDE with singular drift, obtained by (1.1) with $F(u)b$ replaced by $\mathcal{B}$. Namely, for any $\lambda \geq 0$, we consider the backward Cauchy problem

$$\begin{cases} \mathcal{K}u + \langle \mathcal{B}, \nabla_v \rangle u = \lambda u + g, & t \in (0,T), \\ u_T = \ell, \end{cases} \tag{4.1}$$

with $\mathcal{K}$ being the backward Kolmogorov operator associated to the Markovian SDE (2.11), defined in (1.9). Throughout this section we let Assumption 1.1 be in force, and we fix $g, \mathcal{B} \in L_T^\infty\mathcal{C}_B^{-\beta}$ and $\ell \in \mathcal{C}_B^{1+\beta+\varepsilon}$, with $T > 0$, $\beta \in (0, 1/2)$ and $\varepsilon \in (0, 1-2\beta)$.

4.1 Well-posedness of the PDE

Definition 4.1. We say that a function $u \in C_T\mathcal{C}_b \cap L_T^\infty\mathcal{C}_B^{1+\beta+\varepsilon}$ is a *mild solution* for the backward Kolmogorov Cauchy problem (4.1) if the integral equation

$$u_t = P_{T-t}\ell - \int_t^T P_{s-t}\big(\lambda u_s + g_s - \langle \mathcal{B}_s, \nabla_v \rangle u_s\big) ds, \qquad t \in [0,T], \tag{4.2}$$

is satisfied.

Remark 4.2. If $u \in C_T\mathcal{C}_b \cap L_T^\infty\mathcal{C}_B^{1+\beta+\varepsilon}$, then Schauder's estimate (2.16) yields

$$\big\|P_{s-t}\big(\lambda u_s + g_s - \langle \mathcal{B}_s, \nabla_v \rangle u_s\big)\big\|_{1+\beta+\varepsilon} \leq C(s-t)^{-(\beta+\frac{\varepsilon+1}{2})}\|\lambda u_s + g_s - \langle \mathcal{B}_s, \nabla_v \rangle u_s\|_{-\beta} \leq C(s-t)^{-(\beta+\frac{\varepsilon+1}{2})}, \tag{4.3}$$

where

$$\beta + \frac{\varepsilon+1}{2} < 1$$

in light of $\beta \in (0, 1/2)$ and $\varepsilon \in (0, 1-2\beta)$. This shows that the LHS in (4.3) is integrable as a function of s, in $[t,T]$, and thus Definition 4.1 makes sense. Indeed, the integral in (4.2) can be understood as a pointwise Lebesgue integral, namely as

$$\int_t^T P_{s-t}\big(\lambda u_s + g_s - \langle \mathcal{B}_s, \nabla_v \rangle u_s\big)(x) ds, \qquad x \in \mathbb{R}^N,$$

and the standard triangular inequality yields

$$\|u_t\|_{1+\beta+\varepsilon} \leq \|P_{T-t}\phi\|_{1+\beta+\varepsilon} + \int_t^T \big\|P_{s-t}\big(\lambda u_s + g_s - \langle \mathcal{B}_s, \nabla_v \rangle u_s\big)\big\|_{1+\beta+\varepsilon} ds.$$

Definition 4.3. We say that a function $u \in C_T\mathcal{C}_b \cap L^\infty_T \mathcal{C}^{1+\beta+\varepsilon}_B$ is a *weak solution* to the backward Kolmogorov Cauchy problem (4.1) if it is a distributional solution. Namely, for all $\varphi \in \mathcal{S}$ we have

$$\langle \ell \,|\, \varphi\rangle = \langle u_t |\, \varphi\rangle - \int_t^T \langle u_s |\, \mathcal{A}'\,\varphi\rangle\, ds - \sum_{i=1}^d \int_t^T \big\langle (\mathcal{B}_s)_i (\partial_{v_i} u_s) |\, \varphi\big\rangle ds + \int_t^T \langle \lambda u_s + g_s |\, \varphi\rangle ds, \qquad t \in [0,T],$$

where $\mathcal{A}'$ is the formal adjoint of $\mathcal{A}$ as defined in (2.14).

The proof of the next statement is fully analogous to that of Proposition 3.7, and thus is omitted.

Proposition 4.4. *A function $u \in C_T\mathcal{C}_b \cap L^\infty_T \mathcal{C}^{1+\beta+\varepsilon}_B$ is a weak solution to the backward Kolmogorov Cauchy problem* (4.1) *if and only if it is a mild solution.*

We now state the first main result of this section.

Theorem 4.5. *Let Assumption 1.1 hold, and let $g, \mathcal{B} \in L^\infty_T \mathcal{C}^{-\beta}_B$ and $\ell \in \mathcal{C}^{1+\beta+\varepsilon}_B$, for some $\varepsilon \in (0, 1-2\beta)$. Then there exists a unique mild solution $u \in C_T\mathcal{C}_b \cap L^\infty_T \mathcal{C}^{1+\beta+\varepsilon}_B$ to* (4.1).

Remark 4.6. If $0 < \varepsilon' < \varepsilon$ then clearly $\ell \in \mathcal{C}^{1+\beta+\varepsilon'}_B$ and the unique solution $u \in C_T\mathcal{C}_b \cap L^\infty_T \mathcal{C}^{1+\beta+\varepsilon'}_B$ coincides with the solution $u \in C_T\mathcal{C}_b \cap L^\infty_T \mathcal{C}^{1+\beta+\varepsilon}_B$, by uniqueness in the larger space $C_T\mathcal{C}_b \cap L^\infty_T \mathcal{C}^{1+\beta+\varepsilon'}_B$.

Proof of Theorem 4.5. To prove Theorem 4.5, we follow the same strategy employed in the parabolic case ([38, Theorem 4.7]), namely $d = N$ and $B \equiv 0$. The idea is to find a unique fixed point $u \in L^\infty_T \mathcal{C}^{1+\beta+\varepsilon}_B$ for the solution map

$$I_t(w) := P_{T-t}\ell - \int_t^T P_{s-t}\big(\lambda w_s + g_s - \langle \mathcal{B}_s, \nabla_v \rangle w_s\big) ds, \qquad t \in [0,T],$$

and to show that $u \in C_T\mathcal{C}_b$.

Relying on the anisotropic Schauder estimate of Theorem 2.10, one can proceed exactly as in the proof of [38, Theorem 4.7] and obtain that I is a contraction from the complete metric space $(L^\infty_T \mathcal{C}^{1+\beta+\varepsilon}_B, \|\cdot\|_{\rho,T,1+\beta+\varepsilon})$ onto itself for $\rho > 0$ suitably large, where $\|f\|_{\rho,T,1+\beta+\varepsilon} := \sup_{t\in[0,T]} e^{-\rho(T-t)} \|f_t\|_{1+\beta+\varepsilon}$ is the analogous backward of the norm defined in (3.13). Banach fixed point theorem then yields that there exists a unique fixed point $u \in L^\infty_T \mathcal{C}^{\beta+\varepsilon+1}_B$ for I. Eventually, a straightforward modification of the proof of Lemma 3.14 yields $u \in C_T\mathcal{C}_b$ and completes the proof. □

Then next result shows that, if $\mathcal{B}_t, g_t$ are Hölder-continuous functions, then the mild solution u is a solution also in some classical sense and enjoys additional regularity.

Proposition 4.7. *Let Assumption 1.1 be in force. If $\ell \in \mathcal{C}^{2+\alpha}_B$, $g \in C_T\mathcal{C}_b \cap L^\infty_T \mathcal{C}^\alpha_B$ for $\alpha \in (0,1)$ and $\mathcal{B} \in C_T\mathcal{C}_b \cap L^\infty_T \mathcal{C}^{\alpha+\nu}_B$ for some $\nu > 0$, then the mild solution u to* (4.1) *belongs to $C^{2+\alpha}_{B,T}$ (see Definition 2.15) and is also a strong Lie solution in the sense of [47, Definition 2.6].*

Remark 4.8. In particular, the function u solves the PDE pointwise, everywhere on $[0,T] \times \mathbb{R}^N$, where the term Yu that appears in $\mathcal{K}u$ makes sense as a Lie derivative (see (2.17)).

Proof of Proposition 4.7. By [47, Theorem 2.7 and Remark 2.9] (see also [14]), (4.1) admits a (unique) Lie solution $u \in C^{2+\alpha}_{B,T}$. Furthermore, by [47, (3.1)], the latter verifies equation (4.2), and thus u coincides with the unique mild solution to (4.1). □

4.2 Continuity properties

We recall it is assumed that $g, \mathcal{B} \in L_T^\infty \mathcal{C}_B^{-\beta}$ and $\ell \in \mathcal{C}_B^{1+\beta+\varepsilon}$, with $\beta \in (0, 1/2)$ and $\varepsilon \in (0, 1-2\beta)$. Let $\mathcal{B}^{(n)}, g^{(n)} \in L_T^\infty \mathcal{C}_B^{-\beta}$ and $\ell^{(n)} \in \mathcal{C}_B^{1+\beta+\varepsilon}$ denote any approximations of b, g and ℓ, respectively, such that

$$\|\mathcal{B} - \mathcal{B}^{(n)}\|_{T,-\beta-\eta}, \|g - g^{(n)}\|_{T,-\beta-\eta}, \|\ell - \ell^{(n)}\|_{1+\beta+\varepsilon-\eta} \to 0 \qquad \text{as } n \to \infty,$$

for all $\eta > 0$. An example of how to construct said approximation is given in by (2.19) and Lemma 2.16 ensures the convergence. For any fixed $n \in \mathbb{N}$, we consider the unique mild (or, equivalently, weak) solution $u^{(n)}$ to the approximate backward Kolmogorov Cauchy problem

$$\begin{cases} \mathcal{K} u + \langle \mathcal{B}^{(n)}, \nabla_v \rangle u = \lambda u + g^{(n)}, & t \in (0,T), \\ u_T = \ell^{(n)}. \end{cases} \tag{4.4}$$

Theorem 4.9. *Let Assumption 1.1 be in force and let $g, \mathcal{B} \in L_T^\infty \mathcal{C}_B^{-\beta}$ and $\phi \in \mathcal{C}_B^{1+\beta+\varepsilon}$ for some $\varepsilon \in (0, 1-2\beta)$.*

(a) Let $\lambda \geq 0$. For any $\eta > 0$, we have

$$\|u^{(n)} - u\|_{T,1+\beta+\varepsilon-\eta} \longrightarrow 0, \qquad \text{as } n \to \infty, \tag{4.5}$$

where $u^{(n)}, u \in C_T \mathcal{C}_b \cap L_T^\infty \mathcal{C}_B^{1+\beta+\varepsilon}$ are the unique mild solutions to (4.4) and (4.1), respectively.

(b) Let $\ell \equiv 0$. There exists $\bar{\lambda} > 0$, dependent on T, β, ε, $\|\mathcal{B}\|_{T,-\beta}$ and $\|g\|_{T,-\beta}$ such that for any $\lambda > \bar{\lambda}$ we have

$$\|u^{(n)}\|_{T,1+\beta+\varepsilon} \leq \frac{1}{2}, \qquad n \in \mathbb{N}. \tag{4.6}$$

Remark 4.10. Combining Part (a) and Part (b) of Theorem 4.9, one trivially obtains that (4.6) also holds for the unique solution u in $C_T \mathcal{C}_b \cap L_T^\infty \mathcal{C}_B^{1+\beta+\varepsilon}$ to the singular Cauchy problem (4.1) with $\lambda > \bar{\lambda}$ and $\phi \equiv 0$.

Proof of Theorem 4.9. We first prove Part (a). It is enough to show the claim for $\eta < \varepsilon$. Owing to the anisotropic Schauder's estimate of Lemma 2.10, the proof is a straightforward modification of the one of [38, Lemma 4.17-(ii)] where we substitute $\|u - u^{(n)}\|_{\rho,T,1+\alpha}$ with $\|u - u^{(n)}\|_{\rho,T,1+\beta+\varepsilon-\eta}$, $\|\mathcal{B} - \mathcal{B}^{(n)}\|_{T,-\beta}$ with $\|\mathcal{B} - \mathcal{B}^{(n)}\|_{T,-\beta-\eta}$, $\|g - g^{(n)}\|_{\rho,T,-\beta}$ with $\|g - g^{(n)}\|_{\rho,T,-\beta-\eta}$ and $\|\ell - \ell^{(n)}\|_{1+\alpha}$ with $\|\ell - \ell^{(n)}\|_{1+\beta+\varepsilon-\eta}$.

We now prove Part (b). Let $n \in \mathbb{N}$ be fixed. Proceeding similarly as in the proof of [38, Proposition 4.13-(ii)] one obtains that $u^{(n)}$ is a solution of

$$u_t^{(n)} = -\int_t^T e^{-\lambda(s-t)} P_{s-t}\big(g_s^{(n)} - \langle \mathcal{B}_s^{(n)}, \nabla_v \rangle u_s^{(n)}\big) ds, \qquad t \in [0,T]. \tag{4.7}$$

In fact it is enough to choose $\eta < \min(\varepsilon, 1-2\beta-\varepsilon)$ and denote by C, indistinctly, any positive constant that depends, at most, on T, β and ε. For any $0 \leq t < s \leq T$, applying the Schauder estimate of Theorem 2.10 yields

$$\|P_{s-t}\big(g_s^{(n)} - \langle \mathcal{B}_s^{(n)}, \nabla_v \rangle u_s^{(n)}\big)\|_{1+\beta+\varepsilon} \leq C(s-t)^{-\frac{1+\varepsilon+\eta+2\beta}{2}} \|g_s^{(n)} - \langle \mathcal{B}_s^{(n)}, \nabla_v \rangle u_s^{(n)}\|_{-\beta-\eta}$$

(by Proposition 2.9, since $\varepsilon - \eta > 0$)

$$\begin{aligned} &\leq C(s-t)^{-\frac{1+\varepsilon+\eta+2\beta}{2}} \big(\|g_s^{(n)}\|_{-\beta-\eta} + \|\mathcal{B}_s^{(n)}\|_{-\beta-\eta} \|\nabla_v u_s^{(n)}\|_{\beta+\varepsilon}\big) \\ &\leq C(s-t)^{-\frac{1+\varepsilon+\eta+2\beta}{2}} \big(\|g_s^{(n)}\|_{-\beta} + \|\mathcal{B}_s^{(n)}\|_{-\beta} \|\nabla_v u_s^{(n)}\|_{\beta+\varepsilon}\big) \end{aligned}$$

$$\leq C(s-t)^{-\frac{1+\varepsilon+\eta+2\beta}{2}}\big(\sup_m \|g_s^{(m)}\|_{-\beta} + \sup_m \|\mathcal{B}_s^{(m)}\|_{-\beta}\|u_s^{(n)}\|_{1+\beta+\varepsilon}\big),$$

which, combined with (4.7), yields

$$\begin{aligned}\|u^{(n)}\|_{T,1+\beta+\varepsilon} &\leq C\big(\sup_m \|g^{(m)}\|_{T,-\beta} + \sup_m \|\mathcal{B}^{(m)}\|_{T,-\beta}\|u^{(n)}\|_{T,1+\beta+\varepsilon}\big)\sup_{t\in[0,T]}\int_t^T e^{-\lambda(s-t)}(s-t)^{-\frac{1+\varepsilon+\eta+2\beta}{2}}ds \\ &\leq C\big(\sup_m \|g^{(m)}\|_{T,-\beta} + \sup_m \|\mathcal{B}^{(m)}\|_{T,-\beta}\|u^{(n)}\|_{T,1+\beta+\varepsilon}\big)\int_0^T e^{-\lambda s}s^{-\frac{1+\varepsilon+\eta+2\beta}{2}}ds\end{aligned}$$

Now, as $\eta < 1-2\beta-\varepsilon$, the latter integral is finite and tends to 0 as $\lambda \to +\infty$. This completes the proof. □

Corollary 4.11. *Under the assumptions of Theorem 4.9 we have $u^{(n)} \to u$ in $C_T\mathcal{C}_b$ as $n \to \infty$, and in particular $u^{(n)}$ is equicontinuous in n as a function on $[0,T]$ with values in $\mathcal{C}_b$.*

Proof. It is a straightforward consequence of Theorem 4.9, Part (a). □

Lemma 4.12. *Let $(\mathcal{B}^{(n)})_{n\in\mathbb{N}} \subset L_T^\infty\mathcal{C}_B^{-\beta}$ be such that $\mathcal{B}^{(n)} \to \mathcal{B}$ in $L_T^\infty\mathcal{C}_B^{-\beta-\eta}$ for some $\eta > 0$. Let $u^{(n)}$ be the mild $\mathbb{R}^N$-valued solution in $L_T^\infty\mathcal{C}_B^{1+\beta+\varepsilon}$, for some $\varepsilon \in (0, 1-2\beta)$ not depending on n, to the Cauchy problem*

$$\begin{cases}\mathcal{K}u^{(n)} + \langle\mathcal{B}^{(n)}, \nabla_v\rangle u^{(n)} = \lambda\, u^{(n)} - \begin{pmatrix}\mathcal{B}^{(n)}\\ 0\end{pmatrix}\\ u_T^{(n)} = 0,\end{cases} \tag{4.8}$$

which is meant component by component, where we recall that $\mathcal{B}^{(n)}$ is d-dimensional and the vector 0 here is $(N-d)$-dimensional. Similarly, let u be the unique solution to

$$\begin{cases}\mathcal{K}u + \langle\mathcal{B}, \nabla_v\rangle u = \lambda\, u - \begin{pmatrix}\mathcal{B}\\ 0\end{pmatrix}\\ u_T = 0.\end{cases} \tag{4.9}$$

There exists $\bar\lambda$ such that for all $\lambda > \bar\lambda$ we have

$$\|\nabla_v u^{(n)}\|_\infty \vee \|\nabla_v u\|_\infty \leq \frac{1}{2}, \qquad n \in \mathbb{N}. \tag{4.10}$$

We set

$$\phi_t^{(n)}(z) := z + u_t^{(n)}(z), \quad \phi_t(z) := z + u_t(z), \qquad t \in [0,T],\ z \in \mathbb{R}^N,$$

and notice that they take values in $\mathbb{R}^N$. Then we have the following.

(i) For any fixed $t \in [0,T]$, the function $\phi_t^{(n)}$ is invertible;

(ii) denoting by $\psi^{(n)}$ the inverse of $\phi^{(n)}$, the sequence $(\psi^{(n)})$ is equicontinuous;

(iii) $\phi^{(n)} \to \phi$ uniformly on $[0,T]\times\mathbb{R}^N$;

(iv) $\psi^{(n)} \to \psi$ uniformly on $[0,T]\times\mathbb{R}^N$ where ψ is the inverse of ϕ;

(v) $\sup_{t,\widetilde{x}} |\psi_t^{(n)}(\widetilde{v},\widetilde{x})| \leq c(1+|\widetilde{v}|)$, $\forall\widetilde{v} \in \mathbb{R}^d$ for a universal constant c independent of n.

Proof. By Theorem 4.5 each component of $u^{(n)}$ is well-defined as weak solution of (4.8). Theorem 4.9 and Remark 4.10 guarantee that the sequence of solutions $u^{(n)}$ converges in $L_T^\infty \mathcal{C}_B^{1+\beta+\varepsilon}$ for some $\varepsilon \in (0, 1-2\beta)$ to the unique weak solution u of (4.9), and that $\bar\lambda$ can be chosen such that (4.10) holds.

To prove Item (i), we fix $t \in [0,T]$ and, for notational simplicity, we omit the variable t. We sometimes write $\phi^{(n)}(z) = (\phi_1^{(n)}(z), \phi_2^{(n)}(z))$, $u^{(n)}(z) = (u_1^{(n)}(z), u_2^{(n)}(z))$ and $z = (v,x)$, where the first component of these vectors is d-dimensional and the second is $(N-d)$-dimensional. Due to the definition of the vector $u^{(n)}(z)$ as solution of the PDE above, and by uniqueness of the PDE, we have that $u_2^{(n)}(z) = 0$ and so

$$\phi^{(n)}(z) = (\phi_1^{(n)}(z), \phi_2^{(n)}(z)) = (v + u_1^{(n)}(z), x + u_2^{(n)}(z)) = (v + u_1^{(n)}(z), x).$$

To show that $\phi^{(n)}(v,x)$ is invertible it is equivalent to find a unique solution (v,x) of the system

$$\begin{cases} \widetilde v = u_1^{(n)}(v,x) + v \\ \widetilde x = x, \end{cases} \tag{4.11}$$

for any given $\widetilde z = (\widetilde v, \widetilde x) \in \mathbb{R}^N$. If it exists, we denote the inverse of $\phi^{(n)}$ by $\psi^{(n)} = (\psi_1^{(n)}, \psi_2^{(n)}) \in \mathbb{R}^N$. Clearly $x := \widetilde x$ so the inverse of the second block is simply the identity, $\psi_2^{(n)}(\widetilde v, \widetilde x) = \widetilde x$. As for the first block, by (4.10) we have

$$\left|u_1^{(n)}(v,x) - u_1^{(n)}(v',x)\right| \le \frac{1}{2}|v - v'|, \qquad v, v' \in \mathbb{R}^d,$$

uniformly in (t,x), so the map $\widetilde v - u_1^{(n)}(\cdot, x)$ is a contraction, with fixed point in $\mathbb{R}^d$ denoted by $\psi_1^{(n)}(\widetilde v, \widetilde x)$. It immediately follows that $(v,x) := (\psi_1(\widetilde v, \widetilde x), \psi_2(\widetilde v, \widetilde x))$ solves (4.11) and so $\psi^{(n)}(\widetilde z) = (\psi_1^{(n)}(\widetilde z), \psi_2^{(n)}(\widetilde z))$ is the inverse of $\phi^{(n)}(z)$.

We now verify Item (ii), that is the equicontinuity of $(\psi^{(n)})$; to this aim, it is enough to consider only the first component $\psi_1^{(n)}$ because $\psi_2^{(n)}$ is simply the identity in the $\widetilde x$ variable. For $(\psi_1^{(n)})$ we have

$$|\psi_1^{(n)}(t,\widetilde v,\widetilde x) - \psi_1^{(n)}(t',\widetilde v',\widetilde x')| \le |\psi_1^{(n)}(t,\widetilde v,\widetilde x) - \psi_1^{(n)}(t',\widetilde v,\widetilde x')| + |\psi_1^{(n)}(t',\widetilde v,\widetilde x') - \psi_1^{(n)}(t',\widetilde v',\widetilde x')| =: J_1^{(n)} + J_2^{(n)}. \tag{4.12}$$

For the term $J_1^{(n)}$ we can reduce the problem to the equicontinuity of $(u_1^{(n)})$ in (t,x) uniformly in v as follows: setting

$$v := \psi_1^{(n)}(t,\widetilde v,\widetilde x), \qquad v' := \psi_1^{(n)}(t',\widetilde v,\widetilde x'),$$

and using that $\widetilde x = x$ and $\widetilde x' = x'$ we get

$$\begin{aligned} \widetilde v &= \phi_1^{(n)}(t,v,x) = v + u_1^{(n)}(t,v,x) \\ \widetilde v &= \phi_1(t',v',x') = v' + u_1^{(n)}(t',v',x'). \end{aligned}$$

Then,

$$\begin{aligned} |\psi_1^{(n)}(t,\widetilde v,\widetilde x) - \psi_1^{(n)}(t',\widetilde v,\widetilde x')| &= |v - v'| \\ &= \left|u_1^{(n)}(t,v,x) - u_1^{(n)}(t',v',x')\right| \\ &\le \left|u_1^{(n)}(t,v,x) - u_1^{(n)}(t,v',x)\right| + \left|u_1^{(n)}(t,v',x) - u_1^{(n)}(t',v',x')\right| \\ &\le \|\nabla_v u_1^{(n)}\|_\infty |v - v'| + \left|u_1^{(n)}(t,v',x) - u_1^{(n)}(t',v',x')\right| \\ &\le \frac{1}{2}|\psi_1^{(n)}(t,\widetilde v,\widetilde x) - \psi_1^{(n)}(t',\widetilde v,\widetilde x')| + \left|u_1^{(n)}(t,v',x) - u_1^{(n)}(t',v',x')\right|, \end{aligned}$$

having used Theorem 4.9 Part (b). Moving on the LHS the first term we have

$$\begin{aligned}|\psi_1^{(n)}(t,\widetilde{v},\widetilde{x})-\psi_1^{(n)}(t',\widetilde{v},\widetilde{x}')| &\le 2\big|u_1^{(n)}(t,v',x)-u_1^{(n)}(t',v',x')\big| \\ &\le 2\big|u_1^{(n)}(t,v',x)-u_1^{(n)}(t',v',x)\big|+2\big|u_1^{(n)}(t',v',x)-u_1^{(n)}(t',v',x')\big| \\ &=: J_{11}^{(n)}+J_{12}^{(n)}.\end{aligned} \tag{4.13}$$

The term $J_{11}^{(n)}$ is controlled by Corollary 4.11. For the term $J_{12}^{(n)}$, since $u_1^{(n)}(t,\cdot)\in\mathcal{C}_B^{1+\beta+\varepsilon}\subset\mathcal{C}_B^{\alpha}$ for every $\alpha\in(0,1)$, by Remarks 2.14 and 2.12 there exists some $\nu\in(0,1)$ such that for all $t\in[0,T]$

$$\|u_1^{(n)}(t,\cdot,\cdot)\|_{C^\nu}\le C\|u_1^{(n)}(t,\cdot,\cdot)\|_{\alpha}\le C\|u_1^{(n)}(t,\cdot,\cdot)\|_{1+\beta}\le C\|u_1^{(n)}\|_{T,1+\beta},$$

where the latter bound is uniformly bounded in n thanks to Theorem 4.9, Part (a). Thus, recalling also that $\widetilde{x}=x$ and $\widetilde{x}'=x'$, we obtain $J_{12}^{(n)}\le C\sup_n\|u_1^{(n)}\|_{T,1+\beta}|\widetilde{x}-\widetilde{x}'|^\nu$.
For the term $J_2^{(n)}$ in (4.12) we notice that

$$|\psi_1^{(n)}(t',\widetilde{v},\widetilde{x}')-\psi_1^{(n)}(t',\widetilde{v}',\widetilde{x}')|\le\|\nabla_{\widetilde{v}}\psi_1\|_\infty|\widetilde{v}-\widetilde{v}'|,$$

so it is enough to show that $\nabla_{\widetilde{v}}\psi_1^{(n)}$ is uniformly bounded. As $\sup_n\|\nabla_v u_1^{(n)}\|_\infty\le\frac{1}{2}$ by (4.10), the Jacobian

$$\nabla_v\phi_1^{(n)}=I_d+\nabla_v u_1^{(n)}$$

is nonsingular for every (t,v,x). By the Inverse function Theorem, ψ_1 is of class C^1 in $\widetilde{v}$; then

$$\|\nabla_{\widetilde{v}}\psi_1^{(n)}\|_\infty=\|(\nabla_v\phi_1^{(n)})^{-1}\|_\infty\le 2. \tag{4.14}$$

Item (iii) states uniform convergence of $\phi^{(n)}$ to ϕ, and it follows from the definition of $\phi^{(n)}$ and ϕ and the fact that $u_1^{(n)}\to u_1$ uniformly by Corollary 4.11.

Item (iv) states uniform convergence of $\psi^{(n)}$ to ψ. Taking into account Item (ii) and the fact that $\psi_2^{(n)}(t,\widetilde{v},\widetilde{x})=\widetilde{x}=\psi_2(t,\widetilde{v},\widetilde{x})$, it is enough to show pointwise convergence of $\psi_1^{(n)}(t,\widetilde{v},\widetilde{x})$ to $\psi_1(t,\widetilde{v},\widetilde{x})$. We define $v^{(n)}:=\psi_1^{(n)}(t,\widetilde{v},\widetilde{x})$ we have by the inverse property

$$\widetilde{v}=\phi_1^{(n)}(t,v^{(n)},\widetilde{x})=v^{(n)}+u_1^{(n)}(t,v^{(n)},\widetilde{x}). \tag{4.15}$$

Since $(u_1^{(n)})$ is a uniformly bounded sequence, also $(v^{(n)})$ is a bounded sequence and thus it converges to some $v^{(0)}$ up to a subsequence. Since $(\phi_1^{(n)})$ converges uniformly by Item (iii), taking the limit as $n\to\infty$ in (4.15) we get $\widetilde{v}=\phi_1(t,v^{(0)},\widetilde{x})$, so by the inverse property again $v^{(0)}=\psi_1(t,\widetilde{v},\widetilde{x})$ which concludes the proof.

Item (v) can be seen with similar ideas as in Item (ii). Again, we drop the variable t. Since $\psi^{(n)}(\widetilde{v},\widetilde{x})=(\psi_1^{(n)}(\widetilde{v},\widetilde{x}),\psi_2^{(n)}(\widetilde{v},\widetilde{x}))=(\psi_1^{(n)}(\widetilde{v},\widetilde{x}),\widetilde{x})$, we focus on the linear growth of the first component. From bound (4.13) with $t'=t$, $\widetilde{x}'=0$ we get

$$|\psi_1^{(n)}(\widetilde{v},\widetilde{x})-\psi_1^{(n)}(\widetilde{v},0)|\le 4\sup_n\|u_1^{(n)}\|_{C_T\mathcal{C}_b},$$

and together with (4.14) we have

$$\begin{aligned}|\psi_1^{(n)}(\widetilde{v},\widetilde{x})| &\le |\psi_1^{(n)}(\widetilde{v},\widetilde{x})-\psi_1^{(n)}(\widetilde{v},0)|+|\psi_1^{(n)}(\widetilde{v},0)-\psi_1^{(n)}(0,0)|+|\psi_1^{(n)}(0,0)| \\ &\le 4\sup_n\|u_1^{(n)}\|_{C_T\mathcal{C}_b}+2|\widetilde{v}|+\sup_n|\psi_1^{(n)}(0,0)| \\ &\le c(1+|\widetilde{v}|),\end{aligned}$$

having used the fact that $\sup_n\|u_1^{(n)}\|_{C_T\mathcal{C}_b}$ is bounded by Corollary 4.11 and that $\sup_n|\psi_1^{(n)}(0,0)|$ is finite by Item (iv). □

5 The singular McKean-Vlasov kinetic SDE

In this section we apply the results on the PDEs to define and show well-posedness for the singular McKean-Vlasov kinetic equation (1.1). Here the canonical space will be $C_T(\mathbb{R}^N)$ equipped, as usual, with the Borel sets with respect to the topology induced by the uniform convergence. In the sequel we will also make use of the standard *canonical process* on the canonical space.

5.1 The martingale problem formulation for the linear case

Let μ_0 be a probability law on $\mathbb{R}^N$. Throughout this subsection we assume that Assumption 1.1 holds and $\mathcal{B} \in C_T\mathcal{C}_B^{-\beta}$ with $\beta \in (0, 1/2)$. We consider the linear case of equation (1.1), which is formally

$$\begin{cases} dV_t = \Big(\mathcal{B}_t(Z_t) + \mathcal{B}_0 Z_t\Big)dt + dW_t \\ dX_t = \mathcal{B}_1 Z_t dt, \end{cases} \tag{5.1}$$

with initial condition $Z_0 \sim \mu_0$, where the $\mathbb{R}^N$-valued stochastic process $Z_t := (V_t, X_t)$ is the unknown, and $B := \begin{pmatrix} \mathcal{B}_0 \\ \mathcal{B}_1 \end{pmatrix}$ is given as in (1.8). The term linear refers to the fact that the corresponding Fokker-Planck equation is linear, and is reflected in the SDE by the fact that the coefficients do not depend on the law of the unknown. Clearly (5.1) is only formal, and its meaning is given below in terms of martingale problem.

The backward Kolmogorov operator $\mathcal{L}$ associated to the SDE (5.1) with drift $\mathcal{B}$ is given by

$$\mathcal{L}u = \mathcal{K}u + \langle \mathcal{B}, \nabla_v \rangle u, \tag{5.2}$$

where $\mathcal{K}$ is given in (1.9) and $u \in C^{AC}([0,T]; \mathcal{S}') \cap L_T^\infty \mathcal{C}_B^{1+\beta}$.

Let $g \in C_T\mathcal{C}_b \cap L_T^\infty \mathcal{C}_B^\varepsilon$ and $f_T \in \mathcal{C}_B^{1+\beta+\varepsilon}$ for some $\varepsilon \in (0, 1-2\beta)$. We denote by u the unique weak (or, equivalently, mild) solution of

$$\begin{cases} \mathcal{L}u = g \\ u_T = f_T \end{cases} \tag{5.3}$$

in the space $C_T\mathcal{C}_b \cap L_T^\infty \mathcal{C}_B^{1+\beta+\varepsilon}$. Let Z be a stochastic process on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and set

$$M_t^u = M_t^u(Z) := u_t(Z_t) - u_0(Z_0) - \int_0^t g_s(Z_s)ds. \tag{5.4}$$

Definition 5.1. Let μ_0 be a probability law on $\mathbb{R}^N$ and $t_0 \in [0, T)$. Let also $\mathcal{B} \in C_T\mathcal{C}_B^{-\beta}$ and let the operator $\mathcal{L}$ with drift $\mathcal{B}$ be given in (5.2). We denote by $\mathrm{MP}(\mathcal{B}, \mu_0; t_0)$ the martingale problem formally formulated as SDE (5.1), with initial condition $Z_{t_0} \sim \mu_0$.

We say that a couple $(Z, \mathbb{P})$ defined on some measurable space $(\Omega, \mathcal{F})$ is a *solution to MP*$(\mathcal{B}, \mu_0; t_0)$ if Z is continuous, $Z_{t_0} \sim \mu_0$ and for all $g \in C_T\mathcal{S}$ and all $f_T \in \mathcal{S}$ we have that $(M_t^u - M_{t_0}^u)_{t \in [t_0, T]}$ is a (local) martingale under $\mathbb{P}$ with respect to the canonical filtration of Z, where M^u is given in (5.4) and u is the solution of (5.3). We say that $\mathrm{MP}(\mathcal{B}, \mu_0; t_0)$ *admits uniqueness* if, given any pair of solutions $(Z^1, \mathbb{P}^1)$ and $(Z^2, \mathbb{P}^2)$, then the law of Z^1 under $\mathbb{P}^1$ is the same as the law of Z^2 under $\mathbb{P}^2$.

If $t_0 = 0$ we employ the abbreviation $\mathrm{MP}(\mathcal{B}, \mu_0)$ in place of $\mathrm{MP}(\mathcal{B}, \mu_0; 0)$.

Remark 5.2. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and Z be a stochastic process on $(\Omega, \mathcal{F})$. Let us denote by $\mathbb{P}_Z$ the law of Z under $\mathbb{P}$ and by $\hat{Z}$ the canonical process on the canonical space $C_T\mathbb{R}^N$. The couple

$(Z, \mathbb{P})$ is a solution to MP$(\mathcal{B}, \mu_0)$ if and only if the couple $(\hat{Z}, \mathbb{P}_Z)$ is a solution to MP$(\mathcal{B}, \mu_0)$. This is a direct consequence of the change of variable theorem and the martingale property.

Remark 5.3. (a) The solution to MP$(\mathcal{B}, \mu_0)$ can be equivalently defined as follows. Let us define the domain

$$\mathcal{D}_{\mathcal{L}} := \{u \in C_T\mathcal{C}_b \cap L^\infty_T \mathcal{C}^{1+\beta+\varepsilon}_B \text{ for some } \varepsilon \in (0, 1-2\beta) : u \text{ is a sol. to (5.3) for some } g \in C_T\mathcal{S}, f_T \in \mathcal{S}\}.$$

Then $(Z, \mathbb{P})$ is a solution to MP$(\mathcal{B}, \mu_0)$ if and only if for all $u \in \mathcal{D}_{\mathcal{L}}$, M^u is a local martingale under $\mathbb{P}$.

(b) We expect that the formulation of the martingale problem could be equivalently expressed with the smaller domain

$$\mathcal{D}'_{\mathcal{L}} := \{u \in C_T\mathcal{C}_b \cap L^\infty_T \mathcal{C}^{1+\beta+\varepsilon}_B \text{ for some } \varepsilon \in (0, 1-2\beta) : u \text{ is a sol. to (5.3) for some } g \in C_T\mathcal{S}, f_T = 0\}.$$

Indeed, an element of $\mathcal{D}_{\mathcal{L}}$ can be decomposed into the sum $u^0 + u^1$ where $u^1 \in \mathcal{D}'_{\mathcal{L}}$ and u^0 is a solution of $\mathcal{L}u^0 = 0$, i.e. an $\mathcal{L}$-harmonic function. We expect that $u^0(t, X_t)$ is a martingale under $\mathbb{P}$, where $(X, \mathbb{P})$ is a solution of MP$(\mathcal{B}, \mu_0)$.

Lemma 5.4. *If $(Z, \mathbb{P})$ is a solution to MP$(\mathcal{B}, \mu_0)$ then the process M^u defined in (5.4) is a local martingale even if u is solution to PDE (5.3) with $g \in C_T\mathcal{C}_b \cap L^\infty_T \mathcal{C}^\varepsilon_B$ for some $\varepsilon > 0$ and $f_T \in \mathcal{C}^{1+\beta+\varepsilon}_B$, when both have compact support.*

Proof. By Lemma 2.16 Parts (ii) and (iii) there exist two sequences $(g^{(n)})_{n\in\mathbb{N}} \subset C_T\mathcal{S}$ and $(f^{(n)}_T)_{n\in\mathbb{N}} \subset \mathcal{S}$ such that $\|g^{(n)} - g\|_{T,\varepsilon} \to 0$ and $\|f^{(n)}_T - f_T\|_{1+\beta+\varepsilon} \to 0$ for some $\varepsilon > 0$. Thus by Theorem 4.9 Part (a) there exists $\varepsilon > 0$ such that $\|u^{(n)} - u\|_{T,1+\beta+\varepsilon} \to 0$. Consequently $g^{(n)} \to g$ and $u^{(n)} \to u$ uniformly on compacts and thus $M^{u^{(n)}} \to M^u$ in the sense of uniform convergence in probability (ucp). We conclude that M^u is a local martingale since the space of local martingales is closed under the ucp topology. □

In the next result we compare the notion of solution to the MP given in Definition 5.1 with the classical notion of solution à là Stroock-Varadhan. Suppose that $\mathcal{B} \in C_T\mathcal{C}_b \cap L^\infty_T \mathcal{C}^\varepsilon_B$, for some $\varepsilon > 0$. We recall that $(Z, \mathbb{P})$ is a solution to the Stroock-Varadhan Martingale Problem with respect to $\mathcal{L}_{\mathcal{B}} := \mathcal{K} + \langle \mathcal{B}, \nabla_v \rangle$ meant in the classical sense, with initial condition μ_0 if Z is continuous, $Z_0 \sim \mu_0$ and, for every $u \in C^{1,2}([0,T] \times \mathbb{R}^N)$ (or equivalently $u \in C^\infty_c$), the process

$$u_t(Z_t) - u_0(Z_0) - \int_0^t (\mathcal{L}_{\mathcal{B}} u)_s(Z_s) ds$$

is a (local) martingale, see e.g. [58, Chapter 6].

Proposition 5.5. *Let Assumption 1.1 hold. Let $(Z, \mathbb{P})$ be a couple where Z is a stochastic process with $Z_0 \sim \mu_0$ and $\mathbb{P}$ is a probability measure on a given measurable space $(\Omega, \mathcal{F})$. Let $\mathcal{B} \in C_T\mathcal{C}_b \cap L^\infty_T \mathcal{C}^\varepsilon_B$, for some $\varepsilon > 0$. Then $(Z, \mathbb{P})$ is a solution to MP$(\mathcal{B}, \mu_0)$ if and only if it is a solution to the classical Stroock-Varadhan martingale problem with respect to $\mathcal{L}_{\mathcal{B}}$ with initial condition μ_0.*

Proof. "⇒" We have to show that for any $u \in C^\infty_c$ then M^u is a local martingale. It is clear that $\mathcal{K}u \in C^\infty_c$ and $\langle \mathcal{B}, \nabla_z u \rangle \in C_T\mathcal{C}_b \cap L^\infty_T \mathcal{C}^\varepsilon_B$, hence $\mathcal{L}u \in C_T\mathcal{C}_b \cap L^\infty_T \mathcal{C}^\varepsilon_B$ for some $\varepsilon > 0$. Moreover since u has compact support and is time-homogeneous, also $\mathcal{L}u$ and $u(T, \cdot)$ have compact support. The claim now follows by Lemma 5.4.

"$\Leftarrow$" Let $g \in C_T\mathcal{S} \subset C_T\mathcal{C}_b \cap L_T^\infty \mathcal{C}_B^\varepsilon$ and $f_T \in \mathcal{S} \subset \mathcal{C}_B^{2+\varepsilon}$. By assumption, $\mathcal{B} \in C_T\mathcal{C}_b \cap L_T^\infty \mathcal{C}_B^\varepsilon$, and thus, by Proposition 4.7, the mild solution u to (5.3) belongs to $C_{B,T}^{2+\alpha}$ for any $\alpha < \varepsilon$ and solves pointwise

$$\mathcal{L}u = g \qquad (t,z) \in [0,T] \times \mathbb{R}^N,$$

where the term Yu that appears in $\mathcal{L}u$, through $\mathcal{K}u$, makes sense as a Lie derivative (see (2.17)). Furthermore, by Lemma 2.18, there exists a sequence $(u^{(n)})$ such that each $u^{(n)} \in C^{1,2}([0,T] \times \mathbb{R}^N)$, and

$$\|u - u^{(n)}\|_T, \ \|(g - \mathcal{L}u^{(n)})\mathbf{1}_K\|_T \longrightarrow 0, \qquad \text{as } n \to +\infty,$$

for any compact $K \subset \mathbb{R}^N$. Thus $u^{(n)} \to u$ and $\mathcal{L}u^{(n)} \to g$ uniformly on compacts and so $M^{u^{(n)}} \to M^u$ ucp, which implies that M^u is a local martingale on $[0,T]$. □

Corollary 5.6. *Under the conditions of Proposition 5.5, a solution to MP$(\mathcal{B}, \mu_0)$ always exists.*

Proof. Under the conditions of Proposition 5.5, by [58, Theorem 12.2.3 and Exercise 12.4.1] we know that there exists a solution to the classical Stroock-Varadhan Martingale Problem with deterministic initial condition z at time s. The solution is constructed on the canonical space and the probability is denoted by $\mathbb{P}_{s,z}$. Moreover $(s,z) \mapsto \mathbb{P}_{s,z}$ is measurable. Thus we can superpose these laws to treat any initial condition μ_0 at time 0, by setting $\mathbb{P}(B) := \int_{\mathbb{R}^N} \mathbb{P}_{0,z}(B)\mu_0(dz)$ where B is a Borel set of the canonical space. For more details on the *superposition* argument, see the proof of existence of [39, Theorem 4.5]. Hence as a consequence of Proposition 5.5 a solution to MP$(\mathcal{B}, \mu_0)$ always exists when $\mathcal{B} \in C_T\mathcal{C}_b \cap L_T^\infty \mathcal{C}_B^\varepsilon$, for some $\varepsilon > 0$. □

Lemma 5.7. *Let the hypotheses of Lemma 4.12 hold and $\bar{\lambda}$, $\phi^{(n)}$ be as in Lemma 4.12. Let Z be a stochastic process on some measurable space equipped with a sequence of probability measures $\mathbb{P}^{(n)}$, $n \in \mathbb{N}$, such that the following holds.*

(a) For any $\lambda > \bar{\lambda}$ we have

$$\phi_t^{(n)}(Z_t) - \phi_0^{(n)}(Z_0) = \int_0^t \lambda(\phi_r^{(n)}(Z_r) - Z_r)dr + \int_0^t BZ_r dr + M_t^{(n)}, \tag{5.5}$$

where $M^{(n)}$ is an N-dimensional martingale under $\mathbb{P}^{(n)}$, with respect to the canonical filtration of Z, with the last $N-d$ components identically zero and the first d components, denoted by $\widetilde{M}^{(n)}$, with covariation

$$[\widetilde{M}^{(n)}, \widetilde{M}^{(n)}]_t = \int_0^t [\nabla_v \phi_r^{(n)} (\nabla_v \phi_r^{(n)})^\top](Z_r)dr; \tag{5.6}$$

(b) $Z_0 \sim \mu_0$ under $\mathbb{P}^{(n)}$ for any $n \in \mathbb{N}$.

(c) Z_t under $\mathbb{P}^{(n)}$ has finite 4th moment for any $t \in [0,T]$.

Then the sequence of the laws of Z under $\mathbb{P}^{(n)}$ is tight.

Proof. Let us define the process $Y^{(n)}$ as

$$Y_t^{(n)} := u_t^{(n)}(Z_t) + Z_t = \phi_t^{(n)}(Z_t). \tag{5.7}$$

We first show that the sequence of the laws of $Y^{(n)}$ under $\mathbb{P}^{(n)}$ is tight. According to [42, Theorem 4.10 in Chapter 2] we need to prove that

$$\lim_{\eta \to \infty} \sup_{n \geq 1} \mathbb{P}^{(n)}(|Y_0^{(n)}| > \eta) = 0 \tag{5.8}$$

and that for every $\varepsilon > 0$

$$\lim_{\delta\to 0}\sup_{n\geq 1}\mathbb{P}^{(n)}\Big(\sup_{\substack{s,t\in[0,T]\\|s-t|\leq\delta}}|Y_t^{(n)}-Y_s^{(n)}|>\varepsilon\Big)=0. \tag{5.9}$$

By (5.7) and recalling that $Z_0=(V_0,X_0)$ we have

$$Y_0^{(n)}=\phi_0^{(n)}(Z_0)=\begin{pmatrix}V_0+u_1^{(n)}(0,Z_0)\\ X_0\end{pmatrix},$$

so the first condition (5.8) for tightness gives

$$\begin{aligned}\mathbb{P}^{(n)}(|Y^{(n)}(0)|>\eta)&\leq\mathbb{P}^{(n)}(|X_0|+|V_0|+|u_1^{(n)}(0,Z_0)|>\eta)\\&\leq\mathbb{P}^{(n)}(|X_0|+|V_0|+\sup_n\|u_1^{(n)}\|_\infty>\eta),\end{aligned}$$

having used the fact that $\sup_n\|u_1^{(n)}\|_\infty$ is finite by Theorem 4.9 Part (a). The right hand side of the latter inequality does not depend on n because the initial condition $(V_0,X_0)\sim\mu_0$ is independent of n, so (5.8) is satisfied.

Concerning the second bound (5.9) for tightness, if we can prove that

$$\mathbb{E}^{(n)}[|Y_t^{(n)}-Y_s^{(n)}|^4]\leq C|t-s|^2,\qquad t,s\in[0,T], \tag{5.10}$$

holds for some positive constant C independent of n, then we conclude exactly as in the proof of [39, Lemma 3.12], using Garsia-Rodemich-Rumsey lemma (see [6, Section 3]). It remains to prove (5.10). To this aim, using assumption (c) we know that $Y_t^{(n)}=\phi_t^{(n)}(Z_t)$ has finite 4th moment since $\phi_t^{(n)}$ has linear growth. We next prove that there exists a constant C depending on T such that

$$\mathbb{E}^{(n)}[|Y_t^{(n)}|^4]\leq C,\qquad t\in[0,T]. \tag{5.11}$$

Indeed, using (5.5) and (5.7) we get

$$Y_t^{(n)}=Y_0^{(n)}+\int_0^t\lambda u_r^{(n)}(Z_r)dr+\int_0^t BZ_r dr+M_t^{(n)}. \tag{5.12}$$

Taking the square and then the expectation, and using the results of Lemma 4.12, in particular that $Z_t=\psi_t^{(n)}(\phi_t^{(n)}(Z_t))=\psi_t^{(n)}(Y_t^{(n)})$ and the uniform in n linear growth property (v) of $\psi^{(n)}$, Jensen's inequality, BDG inequality and (5.6), one obtains, for all $t\in[0,T]$,

$$\begin{aligned}\mathbb{E}^{(n)}[|Y_t^{(n)}|^4]&\leq c\,\mathbb{E}^{(n)}[|Y_0^{(n)}|^4]+c\,\mathbb{E}^{(n)}\Big[\Big|\int_0^t\lambda u_r^{(n)}(Z_r)dr\Big|^4\Big]+c\,\mathbb{E}^{(n)}\Big[\Big|\int_0^t BZ_r dr\Big|^4\Big]+c\,\mathbb{E}^{(n)}\Big[\Big|M_t^{(n)}\Big|^4\Big]\\&\leq c\,\mathbb{E}^{(n)}[(|Z_0|+|u_0^{(n)}(Z_0)|)^4]+c\,\lambda^4T^4\|u^{(n)}\|_\infty^4+c\,|B|^4t^3\,\mathbb{E}^{(n)}\Big[\int_0^t(1+|Y_r^{(n)}|)^4dr\Big]\\&\quad+c\,T(1+\|\nabla_v u^{(n)}\|_\infty)^2\end{aligned}$$

(by the assumption on the 4th moment of μ_0)

$$\begin{aligned}&\leq c\,(1+\|u^{(n)}\|_{T,1+\beta+\varepsilon}^4)+c\,\mathbb{E}^{(n)}\Big[\int_0^t|Y_r^{(n)}|^4dr\Big]\\&\leq c(1+\int_0^t\mathbb{E}^{(n)}[|Y_r^{(n)}|^4]dr),\end{aligned}$$

for some $\varepsilon > 0$, having used Theorem 4.9 Part (b) in the last bound. Notice that the constant c is independent of n (but depends on T, B, μ_0, λ). Applying Gronwall's lemma we show that (5.11) holds. We now notice that (5.12) yields

$$Y_t^{(n)} - Y_s^{(n)} = \int_s^t \lambda u_r^{(n)}(Z_r)dr + \int_s^t BZ_r dr + M_t^{(n)} - M_s^{(n)}.$$

With similar computations as above, together with multidimensional BDG inequality (see e.g. [42, Problem 3.29]), we have

$$\mathbb{E}^{(n)}[|Y_t^{(n)} - Y_s^{(n)}|^4] \leq c\,\mathbb{E}^{(n)}\left[\left|\int_s^t |\lambda u_r^{(n)}(Z_r)| + |B\psi^{(n)}(Y_r^{(n)})|dr\right|^4\right] + c\,\mathbb{E}^{(n)}\left[\left|M_t^{(n)} - M_s^{(n)}\right|^4\right]$$

(by Lemma 4.12-(v))

$$\begin{aligned}
&\leq c\,\|u^{(n)}\|_{T,1+\beta+\varepsilon}|t-s|^4 + c\,|t-s|^3\mathbb{E}^{(n)}\left[\int_s^t (1+|Y_r^{(n)}|)^4 dr\right] \\
&\quad + c\,\mathbb{E}^{(n)}\left[(\int_s^t |\nabla_v u_r^{(n)}(Z_r) + \begin{pmatrix} I_d \\ 0 \end{pmatrix}|^2 dr)^2\right] \\
&\leq c\|u^{(n)}\|_{T,1+\beta+\varepsilon}|t-s|^4 + c|t-s|^3\int_s^t \mathbb{E}^{(n)}[(1+|Y_r^{(n)}|)^4]dr + c(1+\|u^{(n)}\|_{T,1+\beta+\varepsilon})^2|t-s|^2 \\
&\leq c\|u^{(n)}\|_{T,1+\beta+\varepsilon}|t-s|^4 + c|t-s|^3 + c(1+\|u^{(n)}\|_{T,1+\beta+\varepsilon})^2|t-s|^2,
\end{aligned}$$

having used (5.11) in the last bound. This clearly implies (5.10), as wanted. To conclude that the laws of Z under $\mathbb{P}^{(n)}$ are tight we observe that $Z = \psi^{(n)}(Y^{(n)})$ and $(\psi^{(n)})$ are equicontinuous and have linear growth by Lemma 4.12 Parts (ii), (v). Consequently, when restricted to any compact set, the family $(\psi^{(n)})$ is equicontinuous and bounded, so from tightness of the laws of $Y^{(n)}$ under $\mathbb{P}^{(n)}$ the claim follows. □

Notice that the assumption on the finiteness of the 4th moments of Z_t in Lemma 5.7 is not optimal, and it could be relaxed up to requiring finite $(2+\epsilon)$-moment. However, this assumption is sufficiently general for us as we will apply this result to $Z_0 \sim \mu_0$ being equal to a Dirac delta distribution, and the finiteness of the moments for $t > 0$ will follow.

Lemma 5.8. *Let $(\mathcal{B}^{(n)})$ be a sequence in $C_T\mathcal{C}_b \cap L_T^\infty \mathcal{C}_B^\gamma$ for some $\gamma > 0$ and such that $\|\mathcal{B}^{(n)} - \mathcal{B}\|_{T,-\beta-\eta} \to 0$ for some $\eta > 0$. Let Z be the canonical process and we denote by $(Z, \mathbb{P}^{(n)})$ the solution to the Stroock-Varadhan martingale problem with respect to $\mathcal{L}_{\mathcal{B}^{(n)}}$ and initial condition μ_0. Assume that μ_0 has finite 4th moment. Then there exists a subsequence of $\mathbb{P}^{(n)}$ that converges in law to some probability measure $\mathbb{P}$ and $(Z, \mathbb{P})$ is a solution to MP$(\mathcal{B}, \mu_0)$.*

Remark 5.9. Notice that the solution $(Z, \mathbb{P}^{(n)})$ exists by Corollary 5.6.

Proof of Lemma 5.8. The solution $(Z, \mathbb{P}^{(n)})$ is also a solution to MP$(\mathcal{B}^{(n)}, \mu_0)$ by Proposition 5.5. For the moment we assume the hypothesis of Lemma 5.7 hold (this will be proven later), which implies that the sequence of $(\mathbb{P}^{(n)})$ is tight. Thus there exists a subsequence, which we denote again with $(\mathbb{P}^{(n)})$, that converges in law to some Borel probability measure $\mathbb{P}$ on $C_T\mathbb{R}^N$. We write

$$\mathbb{P}^{(n)} \Rightarrow \mathbb{P}. \tag{5.13}$$

We now show that the couple $(Z, \mathbb{P})$ solves the MP$(\mathcal{B}, \mu_0)$. Since for any $n \in \mathbb{N}$, $\mathbb{P}^{(n)} \circ Z_0^{-1} = \mu_0$ the same holds for $\mathbb{P} \circ Z_0^{-1}$. To prove the martingale property, we consider $u \in \mathcal{D}_{\mathcal{L}}$ (see Remark 5.3) along with $g \in C_T\mathcal{S} \subset C_T\mathcal{C}_b \cap L_T^\infty \mathcal{C}_B^\varepsilon$ and $f_T \in \mathcal{S} \subset \mathcal{C}_B^{1+\beta+\varepsilon}$, for some $\varepsilon > 0$, such that u is a mild solution to (5.3). We then consider the process

$$M_t^u := u_t(Z_t) - u_0(Z_0) - \int_0^t g_s(Z_s)ds,$$

and show that it is a $\mathbb{P}$-martingale. To this aim, fixed an $n \in \mathbb{N}$, we consider the solution $u^{(n)} \in C_T\mathcal{C}_b \cap L_T^\infty \mathcal{C}_B^{1+\beta+\varepsilon}$ of

$$\begin{cases} \mathcal{K}u^{(n)} + \langle \mathcal{B}^{(n)}, \nabla_v \rangle u^{(n)} = g \\ u_T^{(n)} = f_T. \end{cases} \tag{5.14}$$

Such $u^{(n)}$ exists by Theorem 4.5. By definition of solution of MP$(\mathcal{B}^{(n)}, \mu_0)$ we have that

$$M_t^{u^{(n)}} = u_t^{(n)}(Z_t) - u_0^{(n)}(Z_0) - \int_0^t g_s(Z_s)ds$$

is a $\mathbb{P}^{(n)}$-martingale. Moreover, by assumption $\|\mathcal{B}^{(n)} - \mathcal{B}\|_{T,-\beta-\eta} \xrightarrow{n\to\infty} 0$. By (4.5) in Theorem 4.9 Part (a), it now follows that the sequence $(u^{(n)})$ converges uniformly to u and is uniformly bounded. The latter implies, the $M^{u^{(n)}}$ and M^u are uniformly bounded. By these facts and (5.13) it can be shown easily that M^u is a $\mathbb{P}$-martingale by using the definition of martingale, in particular that $\mathbb{E}^{\mathbb{P}}[\Phi(Z_r, r \le s)(M_t^u - M_s^u)] = 0$ for all bounded continuous functionals Φ on $C_s(\mathbb{R}^N)$ knowing that $\mathbb{E}^{\mathbb{P}^{(n)}}[\Phi(Z_r, r \le s)(M_t^{u^{(n)}} - M_s^{u^{(n)}})] = 0$ for all bounded continuous functionals Φ.

It is left to prove that we can apply Lemma 5.7. Let $u^{(n)}$ and $\phi^{(n)}$ be defined as in Lemma 4.12, for some $\bar{\lambda}$ independent of n and defined for the sequence $\mathcal{B}^{(n)}$. The reader should be careful about the notation: $u^{(n)}$ in the rest of the proof should not be confused with the $u^{(n)}$ used above in the current proof, which was solution to (5.14). We would like to apply Itô's formula to $\phi_t^{(n)}(Z_t)$ under $\mathbb{P}^{(n)}$ for fixed n, but $\phi^{(n)}$ does not necessarily belong to $C^{1,2}([0,T] \times \mathbb{R}^N)$, hence we perform a further smoothing procedure and to this aim we momentarily drop the superscript (n), everywhere except in $\mathbb{P}^{(n)}$ and $\mathcal{B}^{(n)}$. Recall that $(Z, \mathbb{P}^{(n)})$ solves the Stroock-Varadhan martingale problem with respect to $\mathcal{L}_{\mathcal{B}^{(n)}}$, therefore there exists a probability space $(\Omega, \mathbb{P}^{(n)})$ and a Brownian motion W and a process $Z = (V, X)$ such that

$$\begin{cases} dV_t = \Big(\mathcal{B}_t^{(n)}(Z_t) + \mathcal{B}_0 Z_t\Big)dt + dW_t \\ dX_t = \mathcal{B}_1 Z_t dt \\ Z_0 \sim \mu_0. \end{cases} \tag{5.15}$$

Indeed, equation (5.15) admits existence in law and furthermore Z under $\mathbb{P}^{(n)}$ has finite 4th moment, hence assumptions (b) and (c) of Lemma 5.7 are satisfied. It remains to prove assumption (a). By Proposition 4.7, $u = u^{(n)}$ is a strong Lie solution of (4.8) in $C_{B,T}^{2+\alpha}$ for any $\alpha \in (0,1)$. Let us denote by $u^{(m)}$ a suitable approximating sequence in $C^{1,2}([0,T] \times \mathbb{R}^N)$, which exists by Lemma 2.18 Part (a). Setting $\phi_t^{(m)}(z) := z + u^{(m)}(z)$ we compute $d\phi_t^{(m)}(Z_t)$ by applying Itô's formula component-wise: we obtain

$$\phi_t^{(m)}(Z_t) - \phi_0^{(m)}(Z_0) = \int_0^t \partial_t \phi_s^{(m)}(Z_s)ds + \int_0^t \nabla \phi_s^{(m)}(Z_s)dZ_s + \frac{1}{2}\sum_{i,j=1}^N \partial_{z_i,z_j}\phi_s^{(m)}(Z_s)d[Z^i, Z^j]_s. \tag{5.16}$$

The second term on the RHS of (5.16) gives

$$\int_0^t \nabla \phi_s^{(m)}(Z_s)dZ_s$$

$$= \int_0^t \langle \begin{pmatrix} \mathcal{B}_s^{(n)}(Z_s) \\ 0 \end{pmatrix}, \nabla \phi_s^{(m)}(Z_s) \rangle ds + \int_0^t \langle BZ_s, \nabla \phi_s^{(m)}(Z_s) \rangle ds + \begin{pmatrix} \int_0^t \nabla_v \phi_s^{(m)}(Z_s) dW_s \\ 0 \end{pmatrix}$$
$$= \int_0^t \langle \mathcal{B}_s^{(n)}(Z_s), \nabla_v \phi_s^{(m)}(Z_s) \rangle ds + \int_0^t \langle BZ_s, \nabla \phi_s^{(m)}(Z_s) \rangle ds + \begin{pmatrix} \int_0^t \nabla_v \phi_s^{(m)}(Z_s) dW_s \\ 0 \end{pmatrix},$$

where the previous identity are to be intended for each component of the vector $\phi^{(m)}$. The last term on the RHS of (5.16) can be written, again component-wise, as

$$\frac{1}{2} \sum_{i,j=1}^{N} \partial_{z_i, z_j} \phi_s^{(m)}(Z_s) d[Z^i, Z^j]_s = \frac{1}{2} \Delta_v \phi_s^{(m)}(Z_s) ds,$$

so (5.16) becomes

$$\phi_t^{(m)}(Z_t) - \phi_0^{(m)}(Z_0) = \int_0^t [\partial_t \phi_s^{(m)}(Z_s) + \langle BZ_s, \nabla \phi_s^{(m)}(Z_s) \rangle + \frac{1}{2} \Delta_v \phi_s^{(m)}(Z_s)] ds$$
$$+ \int_0^t \langle \mathcal{B}_s^{(n)}(Z_s), \nabla_v \phi_s^{(m)}(Z_s) \rangle ds + \begin{pmatrix} \int_0^t \nabla_v \phi_s^{(m)}(Z_s) dW_s \\ 0 \end{pmatrix}.$$

We can rewrite this equality in terms of $u^{(m)}$ as

$$u_t^{(m)}(Z_t) + Z_t - u_0^{(m)}(Z_0) - Z_0 \tag{5.17}$$
$$= \int_0^t [\partial_t u_s^{(m)}(Z_s) + \langle BZ_s, \nabla u_s^{(m)}(Z_s) \rangle + \frac{1}{2} \Delta_v u_s^{(m)}(Z_s)] ds + \int_0^t BZ_s ds$$
$$+ \int_0^t \langle \mathcal{B}_s^{(n)}(Z_s), \nabla_v u_s^{(m)}(Z_s) \rangle ds + \int_0^t \begin{pmatrix} \mathcal{B}_s^{(n)}(Z_s) \\ 0 \end{pmatrix} ds + \begin{pmatrix} \int_0^t \nabla_v u_s^{(m)}(Z_s) dW_s \\ 0 \end{pmatrix} + \begin{pmatrix} W_t \\ 0 \end{pmatrix}$$
$$= \int_0^t [Y u_s^{(m)}(Z_s) + \frac{1}{2} \Delta_v u_s^{(m)}(Z_s)] ds + \int_0^t \langle \mathcal{B}_s^{(n)}(Z_s), \nabla_v u_s^{(m)}(Z_s) \rangle ds$$
$$+ \int_0^t BZ_s ds + \int_0^t \begin{pmatrix} \mathcal{B}_s^{(n)}(Z_s) \\ 0 \end{pmatrix} ds + \begin{pmatrix} \int_0^t \nabla_v u_s^{(m)}(Z_s) dW_s \\ 0 \end{pmatrix} + \begin{pmatrix} W_t \\ 0 \end{pmatrix}.$$

By Lemma 2.18 Part (b) we get that $u^{(m)}, Yu^{(m)}, \Delta_v u^{(m)}$ and $\nabla_v u^{(m)}$ converge uniformly on compact sets to $u, Yu, \Delta_v u$ and $\nabla_v u$, respectively, hence the ucp limit of (5.17) gives

$$u_t(Z_t) + Z_t - u_0(Z_0) - Z_0 = \int_0^t [Y u_s(Z_s) + \frac{1}{2} \Delta_v u_s(Z_s)] ds + \int_0^t \langle \mathcal{B}_s^{(n)}(Z_s), \nabla_v u_s(Z_s) \rangle ds$$
$$+ \int_0^t BZ_s ds + \int_0^t \begin{pmatrix} \mathcal{B}_s^{(n)}(Z_s) \\ 0 \end{pmatrix} ds + \begin{pmatrix} \int_0^t \nabla_v u_s(Z_s) dW_s \\ 0 \end{pmatrix} + \begin{pmatrix} W_t \\ 0 \end{pmatrix}.$$

As u is a strong Lie solution of (4.8), the PDE in the latter is satisfied everywhere on $[0,T] \times \mathbb{R}^N$, and thus

$$u_t(Z_t) + Z_t - u_0(Z_0) - Z_0 = \lambda \int_0^t u_s(Z_s) ds + \int_0^t BZ_s ds + \begin{pmatrix} \int_0^t \nabla_v u_s(Z_s) dW_s \\ 0 \end{pmatrix} + \begin{pmatrix} W_t \\ 0 \end{pmatrix},$$

which can be seen to be equality (5.5), where the martingale M is given by

$$M_t = \begin{pmatrix} \int_0^t \nabla_v \phi_s(Z_s) dW_s \\ 0 \end{pmatrix},$$

having used the fact that $\phi_t(x) = x + u_t(x)$. This concludes the proof that Lemma 5.7 can be applied. $\square$

Proposition 5.10. *Let Assumption 1.1 hold and $\mathcal{B} \in C_T\mathcal{C}_B^{-\beta}$ with $\beta \in (0, 1/2)$. Let μ_0 be a probability law on $\mathbb{R}^N$ with finite 4th moment. Then MP$(\mathcal{B}, \mu_0)$ admits existence of a solution according to Definition 5.1.*

Proof. We define

$$\mathcal{B}_t^{(n)} := \Phi_n * \mathcal{B}_t$$

as in (2.19). By Lemma 2.16, $\mathcal{B}^{(n)} \in C_T\mathcal{C}_b \cap L_T^\infty \mathcal{C}_B^\gamma$ for all $\gamma > 0$, and for any $\eta > 0$

$$\|\mathcal{B}^{(n)} - \mathcal{B}\|_{T,-\beta-\eta} \xrightarrow[n\to+\infty]{} 0.$$

Let Z be the canonical process and we denote by $(Z, \mathbb{P}^{(n)})$ the solution to the Stroock-Varadhan martingale problem with respect to $\mathcal{L}_{\mathcal{B}^{(n)}}$. Lemma 5.8 allows to conclude existence of a solution $(Z, \mathbb{P})$. □

We now prove existence and uniqueness of solutions to MP$(\mathcal{B}, \mu_0)$, for a general distribution law μ_0. Concerning existence, the key tool is the superposition of solutions to MP$(\mathcal{B}, \delta_z)$ for any $z \in \mathbb{R}^N$, with δ_z clearly satisfying the finite 4th moment assumption required to apply Proposition 5.10.

Theorem 5.11. *Let Assumptions 1.1 hold, $\mathcal{B} \in C_T\mathcal{C}_B^{-\beta}$ with $\beta \in (0, 1/2)$, and μ_0 be a generic distribution law on $\mathbb{R}^N$. Then MP$(\mathcal{B}, \mu_0)$ admits existence and uniqueness of a solution according to Definition 5.1.*

Proof. Uniqueness. We want to apply Theorem A.2 from the Appendix, so we need to verify that Property P is satisfied for MP$(\mathcal{B}, \mu_0; t_0)$, for any initial condition μ_0 and any $t_0 \in [0, T)$ (see Definition A.1 in Appendix A for the definition of Property P). Without restriction of generality we set $t_0 = 0$. Assume that $(Z^1, \mathbb{P}^1)$ and $(Z^2, \mathbb{P}^2)$ are two solutions with initial condition μ_0. Without loss of generality, we can choose $Z^1 = Z^2 = Z$ where Z is the canonical process. Let now $g \in C_T\mathcal{S}$ and u be the unique solution of $\mathcal{L}u = g$ with terminal condition $u_T = 0$. By the definition of solution of MP$(\mathcal{B}, \mu_0)$

$$\mathbb{E}^{\mathbb{P}^i}\left[\underbrace{u_T(Z_T)}_{=0} - u_0(Z_0) - \int_0^T g_s(Z_s)ds\right] = 0, \qquad \text{for } i = 1, 2,$$

and $\mathbb{P}^1 \circ Z_0^{-1} = \mathbb{P}^2 \circ Z_0^{-1} = \mu_0$. As a result,

$$\mathbb{E}^{\mathbb{P}^1}\left[\int_0^T g_s(Z_s)ds\right] = \mathbb{E}^{\mathbb{P}^2}\left[\int_0^T g_s(Z_s)ds\right].$$

By arbitrariness of $g \in C_T\mathcal{S}$ then $\mathbb{P}^1$ and $\mathbb{P}^2$ have the same marginals for almost all $s \in [0, T]$, and since Z is continuous they have the same marginals for all $s \in [0, T]$, that is Property P is satisfied for MP$(\mathcal{B}, \mu_0)$, for any initial condition μ_0.

Existence. Given any $z \in \mathbb{R}^N$, by Proposition 5.10 there exists a solution $(Z, \mathbb{P}_z)$ to MP$(\mathcal{B}, \delta_z)$ since δ_z has finite 4th moment. As the solution is unique, the map $z \mapsto \mathbb{P}_z$ is measurable from $\mathbb{R}^N$ to the space of probability measures. This can be seen with the same argument as in [58, Exercise 6.7.4]. Given an arbitrary probability law μ_0, we define $\mathbb{P}$ for a Borel set B of the canonical space, by $\mathbb{P}(B) := \int_{\mathbb{R}^N} \mathbb{P}_z(B)d\mu_0(z)$. The couple $(Z, \mathbb{P})$ is a solution to MP$(\mathcal{B}, \mu_0)$, and we recall that the details of this arguments can be seen in the proof of existence of [39, Theorem 4.5], using the current definition of martingale problem. □

The next proposition is the extension of the convergence result stated in Lemma 5.8, when removing the assumption on the finiteness of the 4th moment of the initial condition.

Proposition 5.12. *Let μ_0 be any probability law on $\mathbb{R}^N$. Let $\mathcal{B}^{(n)}, \mathcal{B}$ be as in Lemma 5.8. Let $(Z, \mathbb{P}^{(n)})$ be a solution of the Stroock-Varadhan martingale problem with respect to $\mathcal{L}_{\mathcal{B}^{(n)}}$ with initial condition μ_0. Then $\mathbb{P}^{(n)}$ converges in law to $\mathbb{P}$, where $(Z, \mathbb{P})$ is solution to MP$(\mathcal{B}, \mu_0)$.*

Proof. For every n we proceed to the *disintegration* of $\mathbb{P}^{(n)}$ as in [39, Fact 2 in the proof of Theorem 4.5]. We write

$$\mathbb{P}^{(n)} = \int \mathbb{P}_z^{(n)} d\mu_0(z). \tag{5.18}$$

We show that $\mathbb{P}_z^{(n)}$ solves the Stroock-Varadhan martingale problem with respect to $\mathcal{L}_{\mathcal{B}^{(n)}}$ and initial condition δ_z; this can be done as in Fact 2 of previously mentioned [39, Theorem 4.5]. By the uniqueness of the Stroock-Varadhan martingale problem, $\mathbb{P}_z^{(n)}$ in (5.18) is the unique solution of that problem. Since $\mu_0 = \delta_z$ has finite 4-th moment, we apply Lemma 5.8 to show that for all $z \in \mathbb{R}^N$ we have

$$\mathbb{P}_z^{(n)} \Rightarrow \mathbb{Q}_z,$$

where $\mathbb{Q}_z$ solves the MP$(\mathcal{B}, \delta_z)$. The map $z \mapsto \mathbb{Q}_z$ is measurable thanks to the uniqueness by Theorem 5.11 and again an argument as [58, Exercise 6.7.4]. By superposition, $\mathbb{Q} := \int_{\mathbb{R}^N} \mathbb{Q}_z d\mu_0(z)$ solves the MP$(\mathcal{B}, \mu_0)$ which equals $\mathbb{P}$ by uniqueness, see again Theorem 5.11. It remains to show that $\mathbb{P}^{(n)} \Rightarrow \mathbb{P}$. Let $F : C_T\mathbb{R}^N \to \mathbb{R}$ be bounded and continuous. We get by Fubini

$$\mathbb{E}^{\mathbb{P}^{(n)}}(F(Z)) = \int_{\mathbb{R}^N} \mathbb{E}^{\mathbb{P}_z^{(n)}}(F(Z)) d\mu_0(z). \tag{5.19}$$

Using the fact that $\mathbb{P}_z^{(n)} \Rightarrow \mathbb{Q}_z$ for all $z \in \mathbb{R}^N$, taking the limit in (5.19) and using dominated convergence theorem we get

$$\lim_{n\to\infty} \mathbb{E}^{\mathbb{P}^{(n)}}(F(Z)) = \int_{\mathbb{R}^N} \mathbb{E}^{\mathbb{Q}_z}(F(Z)) d\mu_0(z) = \mathbb{E}^{\mathbb{P}}(F(Z)),$$

and the proof is concluded. □

In the next result we see that if the solution Z_t of the singular martingale problem has a density v_t in a suitable space, then the function $t \mapsto v_t$ satisfies the Fokker-Planck equation, as expected.

Proposition 5.13. *Let Assumption 1.1 holds and $\mathcal{B} \in C_T\mathcal{C}_B^{-\beta}$ with $\beta \in (0, 1/2)$. Let μ_0 be a probability measure on $\mathbb{R}^N$ with density $u^0 \in \mathcal{C}_B^{\beta+\varepsilon}$, for some $\varepsilon > 0$. Let $(Z, \mathbb{P})$ be the solution to the MP$(\mathcal{B}, \mu_0)$. Then Z_t admits a law density u_t with $u \in C_T\mathcal{C}_b \cap L_T^\infty \mathcal{C}_B^{\beta+\varepsilon}$. Furthermore, u is a weak solution of the linear Fokker-Planck equation* (3.1) *where $F(u)b = \mathcal{B}$, that is for every $\varphi \in \mathcal{S}$ we have*

$$\langle u_t | \varphi \rangle = \langle u^0 | \varphi \rangle + \int_0^t \langle u_s | \mathcal{A} \varphi \rangle ds + \sum_{i=1}^d \int_0^t \langle u_s \mathcal{B}_i | \partial_{z_i} \varphi \rangle ds, \qquad t \in [0, T],$$

where $\mathcal{A}$ was defined in (1.9).

Proof. Denote by u and $u^{(n)}$ the unique mild, and thus weak, solutions in $C_T\mathcal{C}_b \cap L_T^\infty \mathcal{C}_B^{\beta+\varepsilon}$ to (3.1) with $F(u)b = \mathcal{B}$ and $F(u)b = \mathcal{B}^{(n)}$, respectively, which exist by Theorem 3.8. Here $(\mathcal{B}^{(n)})$ is defined as in (2.19). Recall that, by Theorem 3.15, $u^{(n)}$ converges uniformly to u on $[0, T] \times \mathbb{R}^N$.

Furthermore, once more by Theorem 3.15, $u^{(n)}$ is non-negative and $u^{(n)} \in L^1(\mathbb{R}^N)$. By [26, Theorem 2.6], there exists a solution $(Z^{(n)}, \mathbb{P}^{(n)})$ to the Stroock-Varadhan martingale problem with $\mathcal{L}_{\mathcal{B}^{(n)}}$ and initial condition μ_0, whose law density at time t is exactly $u_t^{(n)}$. Owing to Proposition 5.5, the pair $(Z^{(n)}, \mathbb{P}^{(n)})$ coincides with the unique solution to MP$(\mathcal{B}^{(n)}, \mu_0)$. Therefore, by Theorem 5.11 (uniqueness) and Proposition 5.12, $u_t^{(n)}(z)dz$ converges to the law of Z_t under $\mathbb{P}$, for any $t \in [0, T]$. This implies that u_t is the density of Z_t under $\mathbb{P}$. □

5.2 The McKean-Vlasov case

Throughout this section we assume that Assumption 1.1 and 1.4 hold, and that Assumption 1.5 holds for Φ being F and $\widetilde{F}$. The definition of solution to MP extends easily to the McKean-Vlasov case (1.1) and (1.2), recalled below for ease of reading:

$$\begin{cases} dV_t = \Big(F\big(u_t(Z_t)\big)b_t(Z_t) + \mathcal{B}_0 Z_t\Big)dt + dW_t \\ dX_t = \mathcal{B}_1 Z_t dt \\ \mu_{Z_t}(dz) = u_t(z)dz, \end{cases} \tag{5.20}$$

for some initial condition $Z_0 \sim \mu_0 = u^0(z)dz$.

To formulate (5.20) as a martingale problem we resort to the linear singular martingale problem $\mathrm{MP}(\mathcal{B}, \mu_0)$ introduced in Section 5.1, with $\mathcal{B} = F(u)b$ and $u \in C_T\mathcal{C}_b \cap L^\infty_T \mathcal{C}^{\beta+\varepsilon}_B$. Precisely, we give the following definition.

Definition 5.14. Let the McKean-Vlasov Martingale Problem formally given by (1.1) and (1.2) be denoted by $\mathrm{MKMP}(F, b, u^0)$. We say that $\mathrm{MKMP}(F, b, u^0)$ admits *existence* if there exists a triple $(Z, \mathbb{P}, u)$ where Z is a continuous stochastic process, $\mathbb{P}$ is a probability measure, and $u \in C_T\mathcal{C}_b \cap L^\infty_T \mathcal{C}^{\beta+\varepsilon}_B$ for some $\varepsilon > 0$, such that the law of Z_t under $\mathbb{P}$ is u_t and $(Z, \mathbb{P})$ is a solution to $\mathrm{MP}(\mathcal{B}, \mu_0)$ where $\mathcal{B} = F(u)b$.

We say that $\mathrm{MKMP}(F, b, u^0)$ admits *uniqueness* if given any two solutions $(Z, \mathbb{P}, u)$ and $(\hat{Z}, \hat{\mathbb{P}}, \hat{u})$ such that $u, \hat{u} \in C_T\mathcal{C}_b \cap L^\infty_T \mathcal{C}^{\beta+\varepsilon}_B$ for some $\varepsilon > 0$, we have that $u = \hat{u}$ and the law of Z under $\mathbb{P}$ is the same as the law of $\hat{Z}$ under $\hat{\mathbb{P}}$.

We now prove the main result, which was stated in Theorem 1.6 and is repeated here for completeness.

Theorem 5.15. *Let Assumptions 1.1 and 1.4 hold. Let also Assumption 1.5 hold with $\Phi = F, \widetilde{F}$. Then MKMP(F, b, u^0) admits uniqueness.*

Remark 5.16. Let us denote by u_t the probability density of the solution of $\mathrm{MKMP}(F, b, u^0)$ at time $t \in [0, T]$. We now illustrate the fact that the space-regularity of the solution decreases as β increases, as expected. To this aim, let us suppose that u^0 is a bounded and smooth function. The optimal regularity of the solution u is obtained when regarding u^0 as an element of $\mathcal{C}^{1-\beta-\eta}_B$ for every $\eta > 0$ small enough. Indeed we apply Theorem 3.8 and Remark 3.9 with $\beta + \varepsilon = 1 - \beta - \eta$, and we see that $u_t \in \mathcal{C}^{1-\beta-\eta}_B$, hence the regularity index $1 - \beta - \eta$ decreases as β increases.

Proof. Existence. We construct a solution $(Z, \mathbb{P}, u)$ with limiting arguments as follows. Let $(b^{(n)}) \subset L^\infty_T \mathcal{C}^\alpha_B$, for all $\alpha > 0$, be the sequence defined as in (2.19) so that $b^{(n)} \to b$ in $L^\infty_T \mathcal{C}^{-\beta-\eta}_B$ for some $\eta > 0$. Let $u^{(n)} \in C_T\mathcal{C}_b \cap L^\infty_T \mathcal{C}^{\beta+\varepsilon}_B$ for some $\varepsilon > 0$ be the corresponding mild solution to the non-linear FP Cauchy problem (3.16), which exists by Theorem 3.15, Part (a), and moreover, $u^{(n)}$ is a non-negative and bounded function which defines a non-negative finite measure $\mu^{(n)}_t$ with density $u^{(n)}_t$. By Proposition 3.7, $u^{(n)}$ is also a weak solution and thus a solution in the sense of distributions. Hence we can apply [26, Theorem 2.6] to infer that there exists a solution $(Z^{(n)}, \mathbb{P}^{(n)})$ to the Stroock-Varadhan martingale problem with $\mathcal{L}_{\mathcal{B}^{(n)}}$, where $\mathcal{B}^{(n)} := F(u^{(n)})b^{(n)}$, and initial condition μ_0, whose law density at time t is exactly $u^{(n)}_t$. Without loss of generality, let us consider the solution $(Z, \mathbb{P}^{(n)})$, where Z is the canonical process on the canonical space $C_T\mathbb{R}^N$. By the continuity of F in Hölder spaces (Lemma 3.1 with $\Phi = F$) and by continuity of the pointwise product (Proposition 2.9) we obtain that $\mathcal{B}^{(n)} \to \mathcal{B} := F(u)b$ in $L^\infty_T \mathcal{C}^{-\beta-\eta}_B$ for some $\eta > 0$, where u is the

mild solution to the FP Cauchy problem (3.1). By the continuity results on the PDE (3.16), see Theorem 3.15, Part (b), we have $u^{(n)} \to u$ in $C_T\mathcal{C}_b \cap L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$ for some $\varepsilon > 0$. By Proposition 5.12, we know that $\mathbb{P}^{(n)} \to \mathbb{P}$, with $(Z, \mathbb{P})$ being a solution to MP$(\mathcal{B}, \mu_0)$. Recall that $u^{(n)}_t$ is the law of Z_t under $\mathbb{P}^{(n)}$. Since, as mentioned earlier $u^{(n)} \to u$ in $C_T\mathcal{C}_b \cap L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$ so u_t is the law of Z_t under $\mathbb{P}$ and by Definition 5.14, we also have that $(Z, \mathbb{P}, u)$ is a solution to MKMP(F, b, u^0).

Uniqueness. Let $(Z, \mathbb{P}, u)$ and $(\hat{Z}, \hat{\mathbb{P}}, \hat{u})$ be two solutions to MKMP(F, b, u^0). Then $u, \hat{u} \in C_T\mathcal{C}_b \cap L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$, for some $\varepsilon > 0$, by definition. So, by Proposition 5.13 with $\mathcal{B} = F(u)b$ and $\mathcal{B} = F(\hat{u})b$ respectively, we have that u and $\hat{u}$ are both weak solutions to the nonlinear FP singular Cauchy problem (3.1). By Proposition 3.7 and Theorem 3.8, that Cauchy problem admits uniqueness in $\mathcal{C}_T\mathcal{C}_b \cap L^\infty_T\mathcal{C}^{\beta+\varepsilon}_B$; consequently $u = \hat{u}$. By definition of solution to MKMP(F, b, u^0), this means that both $(Z, \mathbb{P})$ and $(\hat{Z}, \hat{\mathbb{P}})$ are solutions to the same linear singular martingale problem, namely MP$(F(u)b, \mu_0)$, where μ_0 is a law with density u^0. By Theorem 5.11, we have uniqueness for the MP$(F(u)b, \mu_0)$ and hence the law of Z under $\mathbb{P}$ is the same as the law of $\hat{Z}$ under $\hat{\mathbb{P}}$, which concludes the proof. □

A Martingale Problem: Uniqueness

The purpose of this Appendix is to prove uniqueness of the Martingale Problem using Ethier-Kurtz techniques, which allow to show uniqueness of the law knowing uniqueness of the time-marginals.

Throughout this section we assume that Assumption 1.1 holds and $\mathcal{B} \in C_T\mathcal{C}^{-\beta}_B$ with $\beta \in (0, 1/2)$.

Definition A.1. Let μ_0 be a Borel probability measure on $\mathbb{R}^N$, and $t_0 \in [0, T)$. We say that *Property P* holds for MP$(\mathcal{B}, \mu_0; t_0)$ if given $(Z, \mathbb{P})$ solution to the MP$(\mathcal{B}, \mu_0; t_0)$ then the marginal laws are uniquely determined, that is if $(Z^1, \mathbb{P}^1)$ and $(Z^2, \mathbb{P}^2)$ are two solutions then

$$\mathbb{P}^1 \circ (Z^1)^{-1}_t = \mathbb{P}^2 \circ (Z^2)^{-1}_t, \qquad t \in [t_0, T].$$

Theorem A.2. *If Property P holds for MP$(\mathcal{B}, \mu_0; t_0)$ for every initial condition μ_0 and every $t_0 \in [0, T)$, then we have uniqueness of the MP$(\mathcal{B}, \mu_0)$ for every μ_0.*

Proof. By Remark 5.2, it is enough to prove uniqueness on the canonical space $C_T\mathbb{R}^N$, hence we let Z be the canonical process and consider two solutions $(Z, \mathbb{P})$ and $(Z, \bar{\mathbb{P}})$ to MP$(\mathcal{B}, \mu_0)$. It is sufficient to prove that for every $n \in \mathbb{N}$, $0 \le t_0 < t_1 < \ldots < t_n < T$ and any strictly positive bounded Borel functions $h_0, \ldots, h_n$ we have

$$\mathbb{E}^{\mathbb{P}}\left[\prod_{i=0}^{n} h_i(Z_{t_i})\right] = \mathbb{E}^{\bar{\mathbb{P}}}\left[\prod_{i=0}^{n} h_i(\bar{Z}_{t_i})\right]. \tag{A.1}$$

Indeed, the same equality still holds true when $h_i = \mathbf{1}_{F_i}$ for closed sets F_i (see [58, proof of Theorem 1.1.1] for the approximation of an indicator function of a closed set with positive functions), hence the law of $(Z_{t_0}, \ldots, Z_{t_n})$ is identified.

We proceed by induction on n. The base step $n = 0$ holds by Property P. Let $n \ge 1$ and assume (A.1) holds with n replaced by $n - 1$. We can define two new probability measures on $C_T\mathbb{R}^N$

$$\mathbb{Q}(H) := \frac{1}{\mathbb{E}^{\mathbb{P}}\left[\prod_{i=0}^{n-1} h_i(Z_{t_i})\right]} E^{\mathbb{P}}\left[\mathbf{1}_H \prod_{i=0}^{n-1} h_i(Z_{t_i})\right], \qquad \bar{\mathbb{Q}}(H) := \frac{1}{\mathbb{E}^{\bar{\mathbb{P}}}\left[\prod_{i=0}^{n-1} h_i(Z_{t_i})\right]} E^{\bar{\mathbb{P}}}\left[\mathbf{1}_H \prod_{i=0}^{n-1} h_i(Z_{t_i})\right],$$

for every $H \in \sigma(Z_t,\, 0 \le t \le T)$. We show that the law of $Z_{t_{n-1}}$ under $\mathbb{Q}$ is the same as the law of $Z_{t_{n-1}}$ under $\bar{\mathbb{Q}}$. Indeed for any strictly positive bounded Borel function l we have, using induction step (A.1) with n replaced by $n-1$,

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}[l(Z_{t_{n-1}})] &= \frac{1}{\mathbb{E}^{\mathbb{P}}\left[\prod_{i=0}^{n-1} h_i(Z_{t_i})\right]} E^{\mathbb{P}}\left[l(Z_{t_{n-1}})\prod_{i=0}^{n-1} h_i(Z_{t_i})\right] \\ &= \frac{1}{\mathbb{E}^{\bar{\mathbb{P}}}\left[\prod_{i=0}^{n-1} h_i(Z_{t_i})\right]} E^{\bar{\mathbb{P}}}\left[l(Z_{t_{n-1}})\prod_{i=0}^{n-1} h_i(Z_{t_i})\right] = \mathbb{E}^{\bar{\mathbb{Q}}}[l(Z_{t_{n-1}})].\end{aligned}$$

We can thus denote by $\mu_{n-1} := \mathbb{Q}\circ Z_{t_{n-1}}^{-1} = \bar{\mathbb{Q}}\circ Z_{t_{n-1}}^{-1}$. Next we want to show that $(Z_{|_{[t_{n-1},T]}}, \mathbb{Q})$ and $(Z_{|_{[t_{n-1},T]}}, \bar{\mathbb{Q}})$ solve $\mathrm{MP}(\mathcal{B}, \mu_{n-1}; t_{n-1})$. We know that, for any u such that $\mathcal{L}u = g \in C_T\mathcal{S}$ and $u(T,\cdot) \in \mathcal{S}$,

$$M_t^u - M_{t_n}^u = u_t(Z_t) - u_{t_{n-1}}(Z_{t_{n-1}}) - \int_{t_{n-1}}^t g_s(Z_s)ds, \qquad t_{n-1} < t \le T,$$

is a martingale under $\mathbb{P}$ and under $\bar{\mathbb{P}}$. Let $t_{n-1} < r < t \le T$ and $H \in \sigma(Z_s, 0 \le s \le r)$ be arbitrary, then

$$\mathbb{E}^{\mathbb{Q}}\left[(M_t^u - M_r^u)\,\mathbf{1}_H\right] = \frac{1}{\mathbb{E}^{\mathbb{P}}\left[\prod_{i=0}^{n-1} h_i(Z_{t_i})\right]} E^{\mathbb{P}}\left[(M_t^u - M_r^u)\,\mathbf{1}_H \cdot \prod_{i=0}^{n-1} h_i(Z_{t_i})\right] = 0,$$

the latter being 0 since $(Z,\mathbb{P})$ is a solution to $\mathrm{MP}(\mathcal{B},\mu_0)$ and $\mathbf{1}_H \cdot \prod_{i=0}^{n-1} h_i(Z_{t_i})$ is bounded measurable with respect to $\sigma(Z_s, 0 \le s \le r)$. Thus $(M_t^u - M_{t_n}^u)_{t\in[t_{n-1},T]}$ is a $\mathbb{Q}$-martingale. Analogously, it is also a $\bar{\mathbb{Q}}$-martingale. Hence both $(Z_{|_{[t_{n-1},T]}}, \mathbb{Q})$ and $(Z_{|_{[t_{n-1},T]}}, \bar{\mathbb{Q}})$ are solutions to $\mathrm{MP}(\mathcal{B}, \mu_{n-1}; t_{n-1})$. Thus by Property P we know that the marginals at time t_n of the solutions are the same, and thus

$$\mathbb{E}^{\mathbb{Q}}[h_n(Z_{t_n})] = \mathbb{E}^{\bar{\mathbb{Q}}}[h_n(Z_{t_n})].$$

Employing the definitions of $\mathbb{Q}$ and of $\bar{\mathbb{Q}}$, this implies

$$\frac{1}{\mathbb{E}^{\mathbb{P}}\left[\prod_{i=0}^{n-1} h_i(Z_{t_i})\right]}\mathbb{E}^{\mathbb{P}}\left[h_n(Z_{t_n})\prod_{i=0}^{n-1} h_i(Z_{t_i})\right] = \frac{1}{\mathbb{E}^{\bar{\mathbb{P}}}\left[\prod_{i=0}^{n-1} h_i(Z_{t_i})\right]}\mathbb{E}^{\bar{\mathbb{P}}}\left[h_n(Z_{t_n})\prod_{i=0}^{n-1} h_i(Z_{t_i})\right].$$

Finally, by the inductive assumption $\mathbb{E}^{\mathbb{P}}\left[\prod_{i=0}^{n-1} h_i(Z_{t_i})\right] = \mathbb{E}^{\bar{\mathbb{P}}}\left[\prod_{i=0}^{n-1} h_i(Z_{t_i})\right]$, we obtain (A.1) as wanted. □

B Schauder Estimates (proof of Theorem 2.10)

Throughout this section, we will denote by C, indistinctly, any positive constant. In each statement below, C depends on a set of parameters, which will be specified in each case.

We start by recalling some known Gaussian bounds on the the fundamental solution Γ of $\mathcal{K}$ defined in (2.12), and its derivatives.

Notation B.1. Let $\nu = (\nu_1, \dots, \nu_N) \in \mathbb{N}_0^N$ be a multi-index. We define the B-length of ν as

$$[\nu]_B := \sum_{j=0}^{r}(2j+1)\sum_{i=\bar{d}_{j-1}+1}^{\bar{d}_j}\nu_i,$$

where $\bar{d}_j$ are defined in terms of d_j, see (1.12). Moreover, as usual $\partial_x^\nu = \partial_{x_1}^{\nu_1}\cdots\partial_{x_N}^{\nu_N}$.

Lemma B.2. *For any $\nu \in \mathbb{N}_0^N$, there exists $C > 0$, only dependent on B and ν, such that*

$$\int_{\mathbb{R}^N} \big(|\partial_y^\nu \Gamma_t(z-y)| + |\partial_y^\nu \Gamma_t(z-e^{Bt}y)|\big)dz \le Ct^{-\frac{[\nu]_B}{2}}, \qquad t>0,\ y\in\mathbb{R}^N.$$

Proof. For $[\nu]_B \le 4$, it is a particular case of [46, Proposition 6 and Lemma 8]. The case $[\nu]_B > 4$ is a straightforward extension: we omit the details for brevity. □

We now continue by proving a fundamental estimate on the Paley-Littlewood blocks of Γ_t. This extends a very similar result, [64, Lemma 3.1], which is given for the kinetic case of Example 1.3.

Lemma B.3. *For any $l \ge 0$ and $T > 0$, there is a positive constant $C = C(B, l, T)$ such that*

$$\|\Delta_j \Gamma_t\|_{L^1} + \|\Delta_j\left(\Gamma_t \circ e^{tB}\right)\|_{L^1} \le C(t4^j)^{-l}, \qquad t\in(0,T],\ j\in\mathbb{N}_0\cup\{-1\}. \tag{B.1}$$

Note that C above is independent of j and t.

Proof. For brevity, we give a complete proof of the estimate on $\|\Delta_j\Gamma_t\|_{L^1}$; analogous arguments apply to $\|\Delta_j\left(\Gamma_t\circ e^{tB}\right)\|_{L^1}$. It is straightforward to prove that $\|\Delta_{-1}\Gamma_t\|_{L^1}$ is bounded, uniformly with respect to $t\in[0,T]$. Therefore, it is enough to prove (B.1) for $j\in\mathbb{N}_0$.

Fix $t\in(0,T]$ and $j\in\mathbb{N}_0$. We first prove the uniform boundedness of the norms of the blocks. For any $z\in\mathbb{R}^N$ we have

$$\begin{aligned}\Delta_j\Gamma_t(z) &= \big(\check\rho_j * \Gamma_t\big)(z) && \text{(by (2.4))}\\ &= 2^{jQ}\big((\check\rho_0\circ D_{2^j})*\Gamma_t\big)(z) = \int_{\mathbb{R}^N}\check\rho_0(y)\Gamma_t\big(z-2^{-j}.y\big)dy. && \text{(B.2)}\end{aligned}$$

As a result,

$$\begin{aligned}\|\Delta_j\Gamma_t\|_{L^1} &= \int_{\mathbb{R}^N}\Big|\int_{\mathbb{R}^N}\check\rho_0(y)\Gamma_t\left(z-2^{-j}.y\right)dy\Big|dz\\ &\le \int_{\mathbb{R}^N}\big|\check\rho_0(y)\big|\underbrace{\Big(\int_{\mathbb{R}^N}\Gamma_t\big(z-2^{-j}.y\big)dz\Big)}_{=1}dy = \|\check\rho_0\|_{L^1}.\end{aligned}$$

Therefore, to prove (B.1), it is not restrictive to assume that $t4^j\ge 1$. For any $n\in\mathbb{N}$ we set the operators

$$\Delta_i^n := (-1)^n\sum_{k=\bar d_{i-1}+1}^{\bar d_i}\partial_{y_k}^{2n}, \qquad i=0,\cdots,r.$$

Set also $p=\prod_{i=0}^r(2i+1)$ and define

$$\Lambda^m := \Delta_0^{pm} + \Delta_1^{\frac{p}{3}m}\cdots+\Delta_r^{\frac{p}{2r+1}m}, \qquad m\in\mathbb{N}.$$

By the property of the Fourier transform, the inverse of Λ^m is defined for those $f\in\mathcal{S}$ such that $\operatorname{supp}\widehat f$ does not contain the origin; in this case such operator acts as

$$\Lambda^{-m}f = \mathcal{F}^{-1}\left(\frac{\widehat f(\xi)}{\sum_{i=0}^r\sum_{k=\bar d_{i-1}+1}^{\bar d_i}|\xi_k|^{2m\frac{p}{2i+1}}}\right) =: \mathcal{F}^{-1}\left(\frac{\widehat f(\xi)}{S_m(\xi)}\right).$$

Now note that, by the chain rule, we have

$$\nabla_y\Gamma_t\big(z-2^{-j}.y\big) = 2^{-j}.\nabla_\zeta\Gamma_t\big(z-\zeta\big)|_{\zeta=2^{-j}.y}, \qquad y,z\in\mathbb{R}^N,$$

It is important to see that the i-th block of the gradient ∇_ζ is multiplied by $2^{-j(2i+1)}$ every time that $\Gamma_t(z-2^{-j}.y)$ is differentiated with respect to a variable of the i-th block. By this fact it is not hard to see that

$$\Lambda^m_y\big(\Gamma_t(z-2^{-j}.y)\big) = 4^{-jpm}\Lambda^m_\zeta\big(\Gamma_t(z-\zeta)\big)\big|_{\zeta=2^{-j}.y}, \qquad y,z\in\mathbb{R}^N, \tag{B.3}$$

where the indeces y and ζ in Λ^m_y and Λ^m_ζ, respectively, denote the variables the operator Λ^m acts on.

We can now conclude the proof. By (B.2) and Plancherel's Theorem, we obtain

$$\Delta_j\Gamma_t(z) = \int_{\mathbb{R}^N}\rho_0(\xi)\overline{\mathcal{F}_y(\Gamma_t(z-2^{-j}.y))(\xi)}d\xi = \int_{\mathbb{R}^N}\frac{\rho_0(\xi)}{S_m(\xi)}\,S_m(\xi)\overline{\mathcal{F}_y(\Gamma_t(z-2^{-j}.y))}(\xi)d\xi$$

(by applying once more Plancherel's Theorem together with the properties of the Fourier transform)

$$= \int_{\mathbb{R}^N}\big(\Lambda^{-m}\check\rho_0\big)(y)\,\Lambda^m_y\big(\Gamma_t(z-2^{-j}.y)\big)\,dy,$$

and thus, by (B.3), we have

$$\|\Delta_j\Gamma_t\|_{L^1} \le 4^{-jpm}\int_{\mathbb{R}^N}\int_{\mathbb{R}^N}\big|\big(\Lambda^{-m}\check\rho_0\big)(y)\big|\times\big|\Lambda^m_\zeta\big(\Gamma_t(z-\zeta)\big)|_{\zeta=2^{-j}.y}\big|\,dydz$$

(by Tonelli's Theorem and the estimate of Lemma B.2 N times with $\nu=\nu_i=(0,\dots,0,\frac{2mp}{2j+1},0,\dots,0)\in\mathbb{N}_0^{i-1}\times\mathbb{N}_0\times\mathbb{N}_0^{N-i}=\mathbb{N}_0^N$ for $i=\bar d_{j-1}+1,\dots,\bar d_j$ and $j=0,\dots,r$)

$$\le C\,4^{-jpm}t^{-pm}\int_{\mathbb{R}^N}\big|\big(\Lambda^{-m}\check\rho_0\big)(y)\big|dy.$$

As $m\in\mathbb{N}$ is arbitrary and $t4^j\ge 1$, this concludes the proof. □

The next two lemmas are also needed to prove Theorem 2.10.

Lemma B.4. *For any $T>0$, there exists a positive $C=C(B,T)$ such that*

$$\big|e^{tB^\top}\xi\big|_B \le C\Bigg(\sum_{\substack{k,h=0,\cdots,r\\ k\ge h}}|t|^{\frac{k-h}{2h+1}}|\xi|_B^{(2k+1)/(2h+1)} + \sum_{\substack{k,h=0,\cdots,r\\ k<h}}|\xi|_B^{(2k+1)/(2h+1)}\Bigg), \qquad \xi\in\mathbb{R}^N,\ t\in[-T,T].$$

Proof. Throughout the proof, we will denote by C any positive constant that depends at most on B and T. We first study the block structure of the powers of $B^\top$. Recalling the structural hypotheses (1.8)-(1.11), we have

$$\big((B^\top)^n\big)_{ij}=0, \qquad \mathbf{b}(j)-\mathbf{b}(i)>n,$$

where $\mathbf{b}(i)$ denotes the block of variables corresponding to the index i, i.e.

$$\mathbf{b}(i) := \min\{k : i\le\bar d_k\}, \qquad i=1,\cdots,N.$$

A graphical illustration of this structure is

$$B^\top=\begin{pmatrix} * & * & 0 & \cdots & 0 & 0\\ * & * & * & 0 & \cdots & 0\\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots\\ * & * & * & \ddots & * & 0\\ * & * & * & \cdots & * & *\\ * & * & * & \cdots & * & * \end{pmatrix}, \qquad (B^\top)^2=\begin{pmatrix} * & * & * & 0 & \cdots & 0 & 0\\ * & * & * & * & 0 & \cdots & 0\\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots\\ * & * & * & \ddots & \ddots & * & 0\\ * & * & * & * & \ddots & * & *\\ * & * & * & * & \cdots & * & *\\ * & * & * & * & \cdots & * & * \end{pmatrix},$$

and so on, until

$$(B^\top)^n = \begin{pmatrix} * & * & \cdots & * & * \\ * & * & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ * & * & \cdots & * & * \\ * & * & \cdots & * & * \end{pmatrix}, \qquad n \geq r-1,$$

with the blocks in the position k, h representing $(d_k \times d_h)$ matrices. As a result, for any $n \in \mathbb{N}$ and $\xi = (\xi_0, \cdots, \xi_r) \in \mathbb{R}^N$, we obtain

$$|t^n (B^\top)^n \xi|_B \leq \sum_{h=0}^{r} \left(C^n \, |t|^n \sum_{k=0}^{(h+n)\wedge r} |\xi_k| \right)^{\frac{1}{2h+1}}$$

(since $|\xi_k| \leq |\xi|_B^{2k+1}$)

$$\leq C^n \sum_{h=0}^{r} |t|^{n/(2h+1)} \sum_{k=0}^{(h+n)\wedge r} |\xi|_B^{(2k+1)/(2h+1)}.$$

For all the pairs of indices (h, k) in the sums above we have

$$n \geq k - h,$$

and thus

$$|t|^{n/(2h+1)} \leq C^n |t|^{(k-h)/(2h+1)}, \qquad t \in [-T, T].$$

This and $|t|^{n/(2h+1)} \leq C^n$ yield

$$|t^n (B^\top)^n \xi|_B \leq C^n \Bigg(\sum_{\substack{k,h=0,\cdots,r \\ k \geq h}} |t|^{\frac{k-h}{2h+1}} |\xi|_B^{(2k+1)/(2h+1)} + \sum_{\substack{k,h=0,\cdots,r \\ k < h}} |\xi|_B^{(2k+1)/(2h+1)} \Bigg)$$

and concludes the proof. □

Lemma B.5. *For any $T > 0$ we have*

$$\left\{ i \in \mathbb{N}_0 \cup \{-1\} \,|\, \mathrm{supp}\rho_j \cap \mathrm{supp}\big(\rho_i \circ e^{tB^\top}\big) \neq \emptyset \right\} \subseteq \Theta_j^t, \qquad t \in [-T, T], \quad j \in \mathbb{N}_0 \cup \{-1\}, \tag{B.4}$$

where

$$\Theta_j^t := \left\{ i \in \mathbb{N}_0 \cup \{-1\} \,\middle|\, 2^i \leq C \sum_{\substack{k,h=0,\cdots,r \\ k \geq h}} |t|^{\frac{k-h}{2h+1}} 2^{\frac{j(2k+1)}{2h+1}},\ 2^j \leq C \sum_{\substack{k,h=0,\cdots,r \\ k \geq h}} |t|^{\frac{k-h}{2h+1}} 2^{\frac{i(2k+1)}{2h+1}} \right\}, \quad j \in \mathbb{N}_0, \tag{B.5}$$

$$\Theta_{-1}^t := \left\{ i \in \mathbb{N}_0 \cup \{-1\} \,\middle|\, 2^i \leq C \sum_{\substack{k,h=0,\cdots,r \\ k \geq h}} |t|^{\frac{k-h}{2h+1}} \right\},$$

with $C = C(B, T)$ positive constant.

Furthermore, for any $\gamma \in \mathbb{R} \setminus \{0\}$, there exists a constant $C = C(B, T, \gamma) > 0$ such that

$$\sum_{i \in \Theta_j^t} 2^{-\gamma i} \leq C \, 2^{-j\gamma} \sum_{\substack{k,h=0,\cdots,r \\ k \geq h}} (|t| 4^j)^{|\gamma|(k-h)/(2h+1)}, \qquad t \in [-T, T], \quad j \in \mathbb{N}_0, \tag{B.6}$$

$$\sum_{i \in \Theta_{-1}^t} 2^{-\gamma i} \leq C, \qquad t \in [-T, T]. \tag{B.7}$$

Proof. Fix $j \in \mathbb{N}_0$. For any $i \in \mathbb{N}_0 \cup \{-1\}$, the second property in (2.2) and the property (2.3) of the partition yield

$$\xi \in \mathrm{supp}\rho_j \cap \mathrm{supp}\big(\rho_i \circ e^{tB^\top}\big) \Longrightarrow \begin{cases} |\xi|_B \le 2^{j+1} \\ |e^{tB^\top}\xi|_B \ge 2^{i-1} - \frac{1}{4} \end{cases},$$

and also

$$\zeta \in \mathrm{supp}\big(\rho_j \circ e^{-tB^\top}\big) \cap \mathrm{supp}\rho_i \Longrightarrow \begin{cases} |e^{-tB^\top}\zeta|_B \ge 2^{j-1} \\ |\zeta|_B \le 2^{i+1} \end{cases}.$$

By applying Lemma B.4 to the first system of inequalities, we have

$$\begin{aligned} 2^{i-1} - \frac{1}{4} \le |e^{tB^\top}\xi|_B &\le C\bigg(\sum_{\substack{k,h=0,\cdots,r \\ k\ge h}} |t|^{\frac{k-h}{2h+1}} |\xi|_B^{(2k+1)/(2h+1)} + \sum_{\substack{k,h=0,\cdots,r \\ k<h}} |\xi|_B^{(2k+1)/(2h+1)} \bigg) \\ &\le C\bigg(\sum_{\substack{k,h=0,\cdots,r \\ k\ge h}} |t|^{\frac{k-h}{2h+1}} 2^{(j+1)(2k+1)/(2h+1)} + \sum_{\substack{k,h=0,\cdots,r \\ k<h}} 2^{(j+1)(2k+1)/(2h+1)} \bigg) \end{aligned}$$

(since $2^{j+1} \ge 1$ then $k < h \Rightarrow 2^{(j+1)(2k+1)/(2h+1)} \le 2^{j+1}$)

$$\le C \sum_{\substack{k,h=0,\cdots,r \\ k\ge h}} |t|^{\frac{k-h}{2h+1}} 2^{(j+1)(2k+1)/(2h+1)}.$$

Analogously, the second system leads to

$$2^{j-1} \le C \sum_{\substack{k,h=0,\cdots,r \\ k\ge h}} |t|^{\frac{k-h}{2h+1}} 2^{(i+1)(2k+1)/(2h+1)},$$

which, combined with the previous inequality, yields (B.4) for $j \in \mathbb{N}_0$. Furthermore, for any $i \in \mathbb{N}_0 \cup \{-1\}$, by (2.2) we have

$$\xi \in \mathrm{supp}\rho_{-1} \cap \mathrm{supp}\big(\rho_i \circ e^{tB^\top}\big) \Longrightarrow \begin{cases} |\xi|_B \le 1 \\ |e^{tB^\top}\xi|_B \ge 2^{i-1} - \frac{1}{4} \end{cases}.$$

Proceeding like above yields (B.4) for $j = -1$.

Prove now (B.6). We only show it when $\gamma > 0$, the case $\gamma < 0$ being easier. Let $i \in \Theta_t^j$. By the second inequality in (B.5) we have

$$2^{-i} \le C 2^{-j} \sum_{\substack{k,h=0,\cdots,r \\ k\ge h}} |t|^{\frac{k-h}{2h+1}} \underbrace{2^{\frac{i(2k+1)}{2h+1} - i}}_{=4^{i(k-h)/(2h+1)}}. \tag{B.8}$$

Now, if $i \le j$, then (B.8) yields

$$2^{-i} \le C 2^{-j} \sum_{\substack{k,h=0,\cdots,r \\ k\ge h}} |t|^{\frac{k-h}{2h+1}} 4^{j(k-h)/(2h+1)} := D_j.$$

For $i > j$ we have $2^{-i} < 2^{-j}$, and by definition of Θ_j^t we get

$$2^{-i} < 2^{-j} \le D_j,$$

that is $i \ge -\log_2 D_j$ for every $i \in \Theta_j^t$. This latter implies

$$\sum_{i\in\Theta_j^t} 2^{-\gamma i} \le \sum_{i=\lfloor -\log_2 D_j \rfloor}^{\infty} 2^{-\gamma i} = \frac{2^{\lfloor \log_2 D_j \rfloor \gamma}}{1 - 2^{-\gamma}} \le \frac{2^\beta D_j^\gamma}{1 - 2^{-\gamma}},$$

which yields (B.6) for $\gamma > 0$. Finally, (B.7) is trivial. □

We are now in the position to prove Theorem 2.10.

Proof of Theorem 2.10. We fix $t \in (0,T]$ and write P_t and P_t' in a convenient way, suitable to exploit the properties of the Fourier transform. We have

$$\Gamma_t * (f \circ e^{-tB})(z) = \int_{\mathbb{R}^N} \Gamma_t(z-y) f(e^{-tB}y)dy = |\det e^{tB}| \int_{\mathbb{R}^N} \Gamma_t\big(z - e^{tB}y\big) f(y)dy, \qquad f \in \mathcal{S},$$

$$\big((\Gamma_t \circ e^{tB}) * (f \circ e^{tB})\big)(y) = \int_{\mathbb{R}^N} \Gamma_t\big(e^{tB}(z-y)\big) f(e^{tB}z)dz = \frac{1}{|\det e^{tB}|} \int_{\mathbb{R}^N} \Gamma_t\big(z - e^{tB}y\big) f(z)dz, \qquad f \in \mathcal{S},$$

and thus

$$P_t' f = \frac{1}{|\det e^{tB}|}\, \Gamma_t * (f \circ e^{-tB}), \qquad P_t f = |\det e^{tB}|\, (\Gamma_t \circ e^{tB}) * (f \circ e^{tB}), \qquad f \in \mathcal{S},$$

which extends to $f \in \mathcal{S}'$ by duality. As the following steps apply to P_t and similarly to P_t', for the sake of brevity, we prove only the part of (2.16) related to P_t'. For any $j \geq -1$ we have obtained that

$$\Delta_j(P_t' f) = \frac{1}{|\det e^{tB}|}\, \big(\Delta_j \Gamma_t\big) * (f \circ e^{-tB}).$$

Furthermore, by the second part of Remark 2.3, we have

$$f \circ e^{-tB} = \sum_{i=-1}^{+\infty} (\Delta_i f) \circ e^{-tB},$$

and thus

$$\Delta_j(P_t' f) = \frac{1}{|\det e^{tB}|} \sum_{i=-1}^{\infty} \big(\Delta_j \Gamma_t\big) * \big((\Delta_i f) \circ e^{-tB}\big).$$

Each term of this summation is nonzero if and only if

$$\mathcal{F}\Big(\big(\Delta_j \Gamma_t\big) * \big((\Delta_i f) \circ e^{-tB}\big)\Big) = |\det e^{tB}|\, \rho_j \widehat{\Gamma}\left(\rho_i \circ e^{tB^\top}\right)\left(\widehat{f} \circ e^{tB^\top}\right)$$

is non-null. Therefore, Lemma B.5 yields

$$\Delta_j(P_t' f) = \frac{1}{|\det e^{tB}|} \sum_{i \in \Theta_j^t} \big(\Delta_j \Gamma_t\big) * \big((\Delta_i f) \circ e^{-tB}\big),$$

and by Young's inequality we obtain

$$\|\Delta_j(P_t' f)\|_{L^\infty} \leq \frac{1}{|\det e^{tB}|} \|\Delta_j \Gamma_t\|_{L^1} \sum_{i \in \Theta_j^t} \underbrace{\|(\Delta_i f) \circ e^{-tB}\|_{L^\infty}}_{=\|\Delta_i f\|_{L^\infty}} \leq \frac{1}{|\det e^{tB}|} \|\Delta_j \Gamma_t\|_{L^1} \|f\|_\gamma \sum_{i \in \Theta_j^t} 2^{-i\gamma}.$$

Assume now $\gamma \neq 0$ and $j \in \mathbb{N}_0$: (B.6) yields

$$\|\Delta_j(P_t' f)\|_{L^\infty} \leq C \|\Delta_j \Gamma_t\|_{L^1} \|f\|_\gamma\, 2^{-j\gamma} \sum_{\substack{k,h=0,\cdots,r \\ k \geq h}} (t4^j)^{|\gamma|(k-h)/(2h+1)}.$$

Now, for each $0 \leq h \leq k \leq r$, we apply Lemma B.3 with

$$l = \frac{\alpha}{2} + \frac{k-h}{2h+1}|\gamma|$$

and obtain

$$(t4^j)^{|\gamma|(k-h)/(2h+1)} \|\Delta_j \Gamma_t\|_{L^1} \leq C(t4^j)^{-l} (t4^j)^{|\gamma| \frac{k-h}{2h+1}} = C t^{-\frac{\alpha}{2}} 2^{-j\alpha}.$$

This proves

$$\|\Delta_j(P'_t f)\|_{L^\infty} \leq C\, t^{-\frac{\alpha}{2}} 2^{-j(\alpha+\gamma)} \|f\|_\gamma, \qquad j \in \mathbb{N}_0. \tag{B.9}$$

Following analogous arguments, the same estimate can be proved for $j = -1$. We omit the details for brevity. Finally, the case $\gamma = 0$ follows by the interpolation Lemma 2.10 in [33]. Precisely, employing the notation in the latter reference, estimate (B.9) for $\gamma = 0$ can be obtained by setting: $\mathbb{X} = L^\infty(\mathbb{R}^N)$, $\beta_0 = -1$, $\beta_1 = 1$, $T_j = \Delta_j P'_t$, $C_{0,j} = C_{-1} 2^{-j\alpha} t^{-\frac{\alpha}{2}}$ and $C_{1,j} = C_1 2^{-j\alpha} t^{-\frac{\alpha}{2}}$, with C_{-1}, C_1 denoting the constant C appearing in (B.9) for $\gamma = -1, 1$ respectively. □

ACKNOWLEDGEMENTS. The research of F.R. was partially supported by the ANR-22-CE40-0015-01 project (SDAIM). The research of Stefano Pagliarani was partially supported by the INdAM - GNAMPA project CUP_E53C22001930001.

References

[1] R. Alexandre and C. Villani, *On the Landau approximation in plasma physics*, in Annales de l'Institut Henri Poincaré C, Analyse non linéaire, vol. 21, Elsevier, 2004, pp. 61–95.

[2] S. Athreya, O. Butkovsky, and L. Mytnik, *Strong existence and uniqueness for stable stochastic differential equations with distributional drift*, The Annals of Probability, 48 (2020), pp. 178–210.

[3] H. Bahouri, J.-Y. Chemin, and R. Danchin, *Fourier analysis and nonlinear partial differential equations*, vol. 343, Springer, 2011.

[4] V. Barbu and M. Röckner, *Probabilistic representation for solutions to nonlinear Fokker-Planck equations*, SIAM J. Math. Anal., 50 (2018), pp. 4246–4260.

[5] V. Barbu, M. Röckner, and F. Russo, *Probabilistic representation for solutions of an irregular porous media type equation: the irregular degenerate case*, Probab. Theory Related Fields, 151 (2011), pp. 1–43.

[6] M. Barlow and M. Yor, *Semi-martingale inequalities via the Garsia-Rodemich-Rumsey lemma, and applications to local times*, Journal of Functional Analysis, 49 (1982), pp. 198 – 229.

[7] R. F. Bass and Z.-Q. Chen, *Brownian motion with singular drift*, Ann. Probab., 31 (2003), pp. 791–817.

[8] N. Belaribi and F. Russo, *Uniqueness for Fokker-Planck equations with measurable coefficients and applications to the fast diffusion equation*, Electron. J. Probab., 17 (2012), pp. no. 84, 28.

[9] S. Benachour, P. Chassaing, B. Roynette, and P. Vallois, *Processus associés à l'équation des milieux poreux*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 23 (1996), pp. 793–832.

[10] S. Biagi and M. Bramanti, *Schauder estimates for Kolmogorov-Fokker-Planck operators with coefficients measurable in time and Hölder continuous in space*, Journal of Mathematical Analysis and Applications, 533 (2024), p. 127996.

[11] P. Blanchard, M. Röckner, and F. Russo, *Probabilistic representation for solutions of an irregular porous media type equation*, Ann. Probab., 38 (2010), pp. 1870–1900.

[12] J. Bony, *Interaction des singularités pour les équations de Klein-Gordon non linéaires*, Séminaire Équations aux dérivées partielles (Polytechnique) dit aussi" Séminaire Goulaouic-Schwartz", (1981), pp. 1–27.

[13] M. Bossy, J.-F. Jabir, and D. Talay, *On conditional McKean Lagrangian stochastic models*, Probab. Theory Related Fields, 151 (2011), pp. 319–351.

[14] M. Bramanti and S. Polidoro, *Fundamental solutions for Kolmogorov-Fokker-Planck operators with time-depending measurable coefficients*, Math. Eng., 2 (2020), pp. 734–771.

[15] R. Catellier and K. Chouk, *Paracontrolled distributions and the 3-dimensional stochastic quantization equation*, Ann. Probab., 46 (2018), pp. 2621–2679.

[16] P. Chaudru de Raynal, J.-F. Jabir, and S. Menozzi, *Multidimensional stable driven McKean-Vlasov SDEs with distributional interaction kernel: a regularization by noise perspective*, Preprint ArXiv 2205.11866, (2022).

[17] ——, *Multidimensional stable driven McKean-Vlasov SDEs with distributional interaction kernel: critical thresholds and related models*, Preprint ArXiv 2302.09900, (2023).

[18] P. Chaudru de Raynal, S. Menozzi, A. Pesce, and X. Zhang, *Heat kernel and gradient estimates for kinetic SDEs with low regularity coefficients*, Bulletin des Sciences Mathématiques, 183 (2023), p. 103229.

[19] P.-E. Chaudru de Raynal and S. Menozzi, *On Multidimensional stable-driven Stochastic Differential Equations with Besov drift*, Electronic Journal of Probability, 27 (2022), pp. 1–52.

[20] J.-Y. Chemin, *Perfect incompressible fluids. Transl. from the French by Isabelle Gallagher and Dragos Iftimie*, vol. 14 of Oxf. Lect. Ser. Math. Appl., Oxford: Clarendon Press, 1998.

[21] R. Danchin, *Fourier analysis methods for pdes*, Lecture notes, 14 (2005).

[22] T. De Angelis, M. Germain, and E. Issoglio, *A numerical scheme for stochastic differential equations with distributional drift*, Stochastic Processes and Their Applications, 154 (2022), pp. 55–90.

[23] F. Delarue and R. Diel, *Rough paths and 1d SDE with a time dependent distributional drift: application to polymers*, Probab. Theory Related Fields, 165 (2016), pp. 1–63.

[24] M. Di Francesco and A. Pascucci, *On a class of degenerate parabolic equations of Kolmogorov type*, AMRX Appl. Math. Res. Express, (2005), pp. 77–116.

[25] H. Dong and T. Yastrzhembskiy, *Global Schauder estimates for kinetic Kolmogorov-Fokker-Planck equations*, Preprint ArXiv 2209.00769, (2022).

[26] A. Figalli, *Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients*, J. Funct. Anal., 254 (2008), pp. 109–153.

[27] F. Flandoli, E. Issoglio, and F. Russo, *Multidimensional SDEs with distributional coefficients*, T. Am. Math. Soc., 369 (2017), pp. 1665–1688.

[28] F. Flandoli, F. Russo, and J. Wolf, *Some SDEs with distributional drift. I. General calculus*, Osaka J. Math., 40 (2003), pp. 493–542.

[29] ——, *Some SDEs with distributional drift. II. Lyons-Zheng structure, Itô's formula and semimartingale characterization*, Random Oper. Stochastic Equations, 12 (2004), pp. 145–184.

[30] N. Garofalo and G. Tralli, *Hardy–Littlewood–Sobolev inequalities for a class of non-symmetric and non-doubling hypoelliptic semigroups*, Mathematische Annalen, 383 (2022), pp. 1–38.

[31] L. Goudenége, E. Haress, and A. Richard, *Numerical approximation of SDEs with fractional noise and distributional drift*, Preprint ArXiv 2302.11455, (2023).

[32] M. Gubinelli, P. Imkeller, and N. Perkowski, *Paracontrolled distributions and singular PDEs*, in Forum of Mathematics, Pi, vol. 3, Cambridge University Press, 2015.

[33] Z. Hao, X. Zhang, R. Zhu, and X. Zhu, *Singular kinetic equations and applications*, Preprint ArXiv 2108.05042, (2021).

[34] L. Hörmander, *Hypoelliptic second order differential equations*, Acta Mathematica, 119 (1967), pp. 147–171.

[35] ——, *The analysis of linear partial differential operators I: Distribution theory and Fourier analysis*, Springer, 2015.

[36] C. Imbert and C. Mouhot, *The Schauder estimate in kinetic theory with application to a toy nonlinear model*, Annales Henri Lebesgue, 4 (2021), pp. 369–405.

[37] C. Imbert and L. Silvestre, *The Schauder estimate for kinetic integral equations*, Analysis & PDE, 14 (2021), pp. 171–204.

[38] E. Issoglio and F. Russo, *A PDE with drift of negative Besov index and linear growth solutions*, Differential and Integral Equations, to appear. Preprint ArXiv 2212.04293, (2022).

[39] E. Issoglio and F. Russo, *SDEs with singular coefficients: the martingale problem view and the stochastic dynamics view*, Preprint Arxiv 2208.10799, (2022).

[40] ——, *McKean SDEs with singular coefficients*, Annales de l'Institut Henri Poincaré. Probabilités et Statistiques., 59 (2023), pp. 1530–1548.

[41] L. Izydorczyk, N. Oudjane, F. Russo, and G. Tessitore, *Fokker-Planck equations with terminal condition and related McKean probabilistic representation*, NoDEA, Nonlinear Differ. Equ. Appl., 29 (2022), p. 41. Id/No 10.

[42] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, Graduate Texts in Mathematics, Springer New York, 1991.

[43] H. Kremp and N. Perkowski, *Multidimensional SDE with distributional drift and Lévy noise*, Bernoulli, 28 (2022), pp. 1757–1783.

[44] E. Lanconelli and S. Polidoro, *On a class of hypoelliptic evolution operators*, vol. 52, 1994, pp. 29–63. Partial differential equations, II (Turin, 1993).

[45] P. Langevin, *Sur la théorie du mouvement brownien*, Compt. Rendus, 146 (1908), pp. 530–533.

[46] G. Lucertini, S. Pagliarani, and A. Pascucci, *Optimal regularity for degenerate kolmogorov equations in non-divergence form with rough-in-time coefficients*, Journal of Evolution Equations, 23 (2023), p. 69.

[47] ——, *Optimal Schauder estimates for kinetic Kolmogorov equations with time measurable coefficients*, Preprint ArXiv 2304.13392, (2023).

[48] A. Lunardi, *Schauder estimates for a class of degenerate elliptic and parabolic operators with unbounded coefficients in* $\mathbb{R}^n$, Annali della Scuola Normale Superiore di Pisa-Classe di Scienze, 24 (1997), pp. 133–164.

[49] S. Menozzi, *Parametrix techniques and martingale problems for some degenerate Kolmogorov equations*, Electron. Commun. Probab., 16 (2011), pp. 234–250.

[50] S. Menozzi, *Martingale problems for some degenerate Kolmogorov equations*, Stochastic Processes and their Applications, 128 (2018), pp. 756–802.

[51] K. Oelschläger, *A law of large numbers for moderately interacting diffusion processes*, Z. Wahrscheinlichkeitstheorie verw Gebiete, 69 (1985), pp. 279–322.

[52] S. Pagliarani, A. Pascucci, and M. Pignotti, *Intrinsic Taylor formula for Kolmogorov-type homogeneous groups*, Journal of Mathematical Analysis and Applications, 435 (2016), pp. 1054–1087.

[53] A. Pascucci, *PDE and martingale methods in option pricing*, vol. 2 of Bocconi & Springer Series, Springer, Milan; Bocconi University Press, Milan, 2011.

[54] A. Pascucci and A. Pesce, *Sobolev embeddings for kinetic Fokker-Planck equations*, Preprint ArXiv 2209.05124, (2022).

[55] N. Perkowski and W. van Zuijlen, *Quantitative heat-kernel estimates for diffusions with distributional drift*, Potential Analysis, online first (2022).

[56] S. Polidoro, *On a class of ultraparabolic operators of Kolmogorov-Fokker-Planck type*, Matematiche (Catania), 49 (1994), pp. 53–105 (1995).

[57] F. Russo and G. Trutnau, *Some parabolic PDEs whose drift is an irregular random noise in space*, Ann. Probab., 35 (2007), pp. 2213–2262.

[58] D. W. Stroock and S. R. S. Varadhan, *Multidimensional diffusion processes*, vol. 233 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 1979.

[59] A. Sznitman, *Topics in propagation of chaos*, vol. 1464 of Ecole d'été de probabilités de Saint-Flour XIX—1989, Springer, 1989.

[60] H. Triebel, *Theory of function spaces. III*, vol. 84 of Monographs in Mathematics, Birkhäuser Verlag, Basel, 2006.

[61] A. Veretennikov, *On weak existence of solutions of degenerate McKean-Vlasov equations*, Preprint ArXiv 2301.01532, (2023).

[62] J. B. Walsh, *An introduction to stochastic partial differential equations*, in École d'été de probabilités de Saint-Flour, XIV—1984, vol. 1180 of Lecture Notes in Math., Springer, Berlin, 1986, pp. 265–439.

[63] X. Zhang, *Second order McKean-Vlasov SDEs and kinetic Fokker-Planck-Kolmogorov equations*, Preprint ArXiv 2109.01273, (2021).

[64] X. Zhang and X. Zhang, *Cauchy problem of stochastic kinetic equations*, Preprint ArXiv 2103.02267, (2021).