Journal of Theoretical Probability
https://doi.org/10.1007/s10959-024-01325-5

Stochastic Differential Equations with Singular Coefficients: The Martingale Problem View and the Stochastic Dynamics View

Elena Issoglio[1] · Francesco Russo[2]

Received: 30 June 2023 / Revised: 16 February 2024 / Accepted: 4 March 2024

Abstract
We consider stochastic differential equations (SDEs) with (distributional) drift in negative Besov spaces and random initial condition and investigate them from two different viewpoints. In the first part we set up a martingale problem and show its well-posedness. We then prove further properties of the martingale problem, such as continuity with respect to the drift and the link with the Fokker–Planck equation. We also show that the solutions are weak Dirichlet processes for which we evaluate the quadratic variation of the martingale component. In the second part we identify the dynamics of the solution of the martingale problem by describing the proper associated SDE. Under suitable assumptions we show equivalence with the solution to the martingale problem.

Keywords Stochastic differential equations · Distributional drift · Besov spaces · Martingale problem · Weak Dirichlet processes

Mathematics Subject Classification (2020) 60H10 · 60H30 · 60H50

1 Introduction

In this paper we study the stochastic differential equation (SDE)

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \mathrm{d}W_t, \quad X_0 \sim \mu, \tag{1.1}$$

✉ Elena Issoglio
elena.issoglio@unito.it

Francesco Russo
francesco.russo@ensta-paris.fr

1 Dipartimento di Matematica 'G. Peano', Universitá di Torino, Turin, Italy

2 Unité de Mathématiques appliquées, ENSTA Paris, Institut Polytechnique de Paris, Palaiseau, France

where $X_t \in \mathbb{R}^d$, the process (W_t) is a d-dimensional Brownian motion, μ is any probability measure and the drift $b(t, \cdot)$ is an element of a negative Besov space $\mathcal{C}^{(-\beta)+}$, see below for the precise definition. SDE (1.1) is clearly only formal at this stage, because the drift b cannot even be evaluated at the point X_t, and one first needs to define a notion of solution for this kind of SDEs. We tackle this problem from two different viewpoints. In the first part we set up a martingale problem and show its well-posedness. In the second part we identify the dynamics of the solution of the martingale problem.

The first steps in the study of the SDE in dimension 1 (and with a diffusion coefficient σ) were done in [13, 14, 22]. In dimension $d > 1$ we mention the work [12] where the authors introduced the notion of *virtual solution* whose construction depended a priori on a real parameter λ. Also, the setting was slightly different because the function spaces were negative fractional Sobolev spaces $H_q^{-\beta}$ and not Besov spaces. Other authors have studied SDEs with distributional coefficients; afterwards, we mention in particular [1, 5, 7, 29]. The main idea in all these works, which is the same we also develop in the first part of the present paper, is to frame the SDE as a martingale problem; hence, the main goal is to find a domain $\mathcal{D}_{\mathcal{L}}$ that characterises the martingale solution in terms of the quantity

$$f(t, X_t) - f(0, X_0) - \int_0^t \mathcal{L}\{(s, X_s)\mathrm{d}s, \tag{1.2}$$

for all $f \in \mathcal{D}_{\mathcal{L}}$, where $\mathcal{L}$ is the parabolic generator of X formally given by $\mathcal{L}f = \partial_t f + \frac{1}{2}\Delta f + \nabla f\, b$. This is made rigorous using results on the PDE

$$\begin{cases} \mathcal{L}f = g \\ f(T) = f_T, \end{cases}$$

developed in [18].

Our framework in terms of function spaces is slightly different than all the works cited above. In the first part of the article, the only difference is that we allow the initial condition X_0 to be any random variable, and not only a Dirac delta in a point x. Well-posedeness of the PDE $\mathcal{L}\{ = g$ allows to give a proper meaning to the martingale problem. Various regularity results on the PDE together with a transformation of the solution X into the solution Y of a 'standard' (Stroock–Varadhan) martingale problem (see Sect. 3) allow us to show existence and uniqueness of the solution X to the martingale problem, see Theorem 4.5. We also prove other interesting results such as Theorem 4.2 where we show that the law density of the solution X satisfies the Fokker–Planck equation, which is a PDE with negative Besov coefficients. Furthermore, we show in Theorem 4.3 some tightness results for smoothed solutions X^n when the negative Besov coefficients are smoothed.

The main novelty of this paper is the second part, where we study the SDE $X_t = X_0 + \int_0^t b(s, X_s)\mathrm{d}s + W_t$ from a different point of view, in particular we look into the dynamics of the process itself. One natural question to ask, which is well understood in the classical Stroock–Varadhan case where b is a locally bounded function, is the equivalence between the solution to the martingale problem and the solution in law (i.e. weak solution) of the SDE. In the case of SDEs with distributional coefficients, the first challenging problem is to define a suitable notion of solution of the SDE and then to study well-posedness of that equation. To this aim, we start in Sect. 5 by showing that the solution to the martingale problem is a weak Dirichlet process, for which we identify the martingale component in its canonical decomposition, see Proposition 5.11 and Remark 5.12. We then introduce in Sect. 6 our notion of solution for the SDE, involving a 'local time' operator which plays the role of the integral $\int_0^t b(s, X_s)\mathrm{d}s$ and involving weak Dirichlet processes. Under further mild assumptions on b, for example if it has compact support, in Theorem 6.5 we show that a solution to the martingale problem is also a solution to the SDE. In a slightly more restricted framework, in Proposition 6.12 we obtain the converse result, hence providing the equivalence result of SDEs and martingale problems for distributional drifts, see Corollary 6.13. Those results extend [22, Propositions 6.7 and 6.10] stated in dimension 1 and in the case of time-homogeneous coefficients.

A typical example of drift b for which all our results are valid, arises when b is a quenched realisation of an independent noise $\dot{B}_x(\omega)$, which is a generalised random field whose trajectories are the divergence of a $(1-\beta)$-Hölder continuous functions $x \mapsto B_x(\omega)$ for some $\beta \in (0, \frac{1}{2})$, cut with a smooth function with compact support. These models arise when describing the motion of particles propagating in an irregular medium, see [27] and references therein. The class of these noises is large, and in dimension $d = 1$ it includes for instance (bi)fractional, multi-fractional Brownian ones, etc., with Hurst index greater than $\frac{1}{2}$, to be cut so that they have compact support.

A result connected to ours is provided by [6], where the authors study the case when the driving noise is a Lévy α-stable process and the distributional drift lives in a general Besov space $\mathbb{B}^{-\beta}_{p,q}$. In particular, they formulate the martingale problem and a quite different notion of SDE (for which, in $d = 1$ they even study pathwise uniqueness, extending in this way [22, Corollary 5.19], stated for Brownian motion) and prove that a solution to the martingale problem is also a solution to their SDE. However, they do not prove the converse result; hence, they do not have any equivalence.

The paper is organised as follows. In Sect. 2 we introduce the framework in which we work, in particular the various functions spaces appearing in the paper and many useful results from the companion paper [18]. In Sect. 3 we introduce the martingale problem and transform it into a classical equivalent Stroock–Varadhan martingale problem. In Sect. 4 we show existence and uniqueness of a solution to the martingale problem and various other properties. In Sect. 5 we show that the solution to the martingale problem is a weak Dirichlet process and identify its decomposition. In Sect. 6 we introduce the notion of solution to the SDE and show its equivalence to the martingale problem. Finally, in Appendix A we state a useful result on solutions of (classical) PDEs that we use in the paper.

2 Setting and Preliminary Results

2.1 Function Spaces

Let us denote by $C^{1,2}_{buc} := C^{1,2}_{buc}([0,T]\times\mathbb{R}^d)$ the space of all $C^{1,2}$ real functions such that the function and its gradient in x are bounded, and the Hessian matrix and the time-derivative are bounded and uniformly continuous. Let us denote by $C^{1,2}_c := C^{1,2}_c([0,T]\times\mathbb{R}^d)$ the space of $C^{1,2}([0,T]\times\mathbb{R}^d)$ with compact support. Let us denote by $C^{1,2}_b := C^{1,2}_b([0,T]\times\mathbb{R}^d)$ the space of $C^{1,2}$-functions that are bounded with bounded derivatives. We also use the notation $C^{0,1} := C^{0,1}([0,T]\times\mathbb{R}^d)$ to indicate the space of real functions with gradient in x uniformly continuous in (t,x). Let $C^\infty_c := C^\infty_c(\mathbb{R}^d)$ denote the space of all smooth real functions with compact support. We denote by $C_c = C_c(\mathbb{R}^d)$ the space of $\mathbb{R}$-valued continuous functions with compact support. Let $\mathcal{S} = \mathcal{S}(\mathbb{R}^d)$ be the space of real-valued Schwartz functions on $\mathbb{R}^d$ and $\mathcal{S}' = \mathcal{S}'(\mathbb{R}^d)$ the space of Schwartz distributions. The corresponding dual pairing will be denoted by $\langle\cdot,\cdot\rangle$.

For $\gamma\in\mathbb{R}$ we denote by $\mathcal{C}^\gamma = \mathcal{C}^\gamma(\mathbb{R}^d)$ the Besov space (or Hölder-Zygmund space), endowed with its norm $\|\cdot\|_\gamma$. For more details see [2, Section 2.7, pag 99] and also [18], where we recall all useful facts and definitions about these spaces. If $\gamma\in\mathbb{R}^+\setminus\mathbb{N}$ then the space coincides with the classical Hölder space. If $\gamma<0$ then the space includes some Schwartz distributions. We have $\mathcal{C}^\gamma\subset\mathcal{C}^\alpha$ for any $\gamma>\alpha$. Moreover, it holds that $L^\infty\subset\mathcal{C}^0$ (see [17] for a proof in the case of anisotropic Besov spaces). We denote by $C_T\mathcal{C}^\gamma$ the space of continuous functions on $[0,T]$ taking values in $\mathcal{C}^\gamma$, that is $C_T\mathcal{C}^\gamma := C([0,T];\mathcal{C}^\gamma)$. For any given $\gamma\in\mathbb{R}$ we denote by $\mathcal{C}^{\gamma+}$ and $\mathcal{C}^{\gamma-}$ the spaces given by

$$\mathcal{C}^{\gamma+} := \cup_{\alpha>\gamma}\mathcal{C}^\alpha, \qquad \mathcal{C}^{\gamma-} := \cap_{\alpha<\gamma}\mathcal{C}^\alpha.$$

Note that $\mathcal{C}^{\gamma+}$ is an inductive space. We will also use the spaces $C_T\mathcal{C}^{\gamma+} := C([0,T];\mathcal{C}^{\gamma+})$, which is equivalent to the fact that for $f\in C_T\mathcal{C}^{\gamma+}$ there exists $\alpha>\gamma$ such that $f\in C_T\mathcal{C}^\alpha$, see for example [19, Appendix B]. Similarly, we use the space $C_T\mathcal{C}^{\gamma-} := C([0,T];\mathcal{C}^{\gamma-})$, meaning that if $f\in C_T\mathcal{C}^{\gamma-}$ then for any $\alpha<\gamma$ we have $f\in C_T\mathcal{C}^\alpha$. We denote by $\mathcal{C}^\gamma_c = \mathcal{C}^\gamma_c(\mathbb{R}^d)$ the space of elements in $\mathcal{C}^\gamma$ with compact support. Similarly when γ is replaced by $\gamma+$ or $\gamma-$. When defining the domain of the martingale problem, we will work with spaces of functions which are the limit of functions with compact support, so that they are Banach spaces. More precisely, let us denote by $\bar{\mathcal{C}}^\gamma_c = \bar{\mathcal{C}}^\gamma_c(\mathbb{R}^d)$ the space

$$\bar{\mathcal{C}}^\gamma_c := \{f\in\mathcal{C}^\gamma \text{ such that } \exists (f_n)\subset\mathcal{C}^\gamma_c \text{ and } f_n\to f \text{ in } \mathcal{C}^\gamma\}.$$

As above we denote the inductive space and intersection space as

$$\bar{\mathcal{C}}^{\gamma+}_c := \cup_{\alpha>\gamma}\bar{\mathcal{C}}^\alpha_c, \qquad \bar{\mathcal{C}}^{\gamma-}_c := \cap_{\alpha<\gamma}\bar{\mathcal{C}}^\alpha_c.$$

The main reason for introducing this class of subspaces is that $\bar{\mathcal{C}}^{\gamma+}_c$ are separable, as proved in [18, Lemma 5.7], unlike the classical Besov spaces $\mathcal{C}^\gamma$ and $\mathcal{C}^{\gamma+}$ which are not separable. Similarly as above, we use the space $C_T\bar{\mathcal{C}}_c^{\gamma+} := C([0,T];\bar{\mathcal{C}}_c^{\gamma+})$; in particular we observe that if $f \in C_T\bar{\mathcal{C}}_c^{\gamma+}$ then for any $\alpha < \gamma$ we have $f \in C_T\bar{\mathcal{C}}_c^{\alpha}$ by [19, Remark B.1, part (ii)]. Moreover, in [18, Corollary 5.8] we show that $C_T\bar{\mathcal{C}}_c^{\gamma+}$ is separable. Note that if f is continuous and such that $\nabla f \in C_T\mathcal{C}^{0+}$ then $f \in C^{0,1}$.

Note that for all function spaces introduced above we use the same notation to indicate $\mathbb{R}$-valued functions but also $\mathbb{R}^d$- or $\mathbb{R}^{d\times d}$-valued functions. It will be clear from the context which space is needed. When $f : \mathbb{R}^d \to \mathbb{R}^m$ is differentiable, we denote by ∇f the matrix given by $(\nabla f)_{i,j} = \partial_i f_j$. In particular, when $f : \mathbb{R}^d \to \mathbb{R}$ then ∇f is a column vector and we denote the Hessian matrix of f by $\mathrm{Hess}(f)$.

For $\gamma \in (0,1)$ we define space $D\mathcal{C}^\gamma$ as

$$D\mathcal{C}^\gamma := \{h : \mathbb{R}^d \to \mathbb{R} \text{ differentiable function s.t. } \nabla h \in \mathcal{C}^\gamma\},$$

and by $C_T D\mathcal{C}^\gamma := C([0,T]; D\mathcal{C}^\gamma)$. Note that the following inclusion holds $\mathcal{C}^{1+\alpha} \subset D\mathcal{C}^\alpha$. Analogously as for the $\mathcal{C}^{\gamma+}$-spaces, for $\gamma > 0$ we also introduce the spaces

$$D\mathcal{C}^{\gamma+} := \cup_{\alpha>\gamma} D\mathcal{C}^\alpha, \qquad D\mathcal{C}^{\gamma-} := \cap_{\alpha<\gamma} D\mathcal{C}^\alpha.$$

We will also use the spaces $C_T D\mathcal{C}^{\gamma+} := C([0,T]; D\mathcal{C}^{\gamma+})$. For more details on these spaces, see [18, Sect. 3].

2.2 Some Tools and Properties

The following is an important estimate which allows to define the pointwise product between certain distributions and functions, which is based on Bony's estimates. For details see [4] or [16, Sect. 2.1]. Let $f \in \mathcal{C}^\alpha$ and $g \in \mathcal{C}^{-\beta}$ with $\alpha - \beta > 0$ and $\alpha, \beta > 0$. Then the 'pointwise product' fg is well-defined as an element of $\mathcal{C}^{-\beta}$ and there exists a constant $c > 0$ such that

$$\|fg\|_{-\beta} \le c\|f\|_\alpha \|g\|_{-\beta}. \tag{2.1}$$

Remark 2.1 Using (2.1) it is not difficult to see that if $f \in C_T\mathcal{C}^\alpha$ and $g \in C_T\mathcal{C}^{-\beta}$ for $\alpha > \beta > 0$ then the product is also continuous with values in $\mathcal{C}^{-\beta}$, and

$$\|fg\|_{C_T\mathcal{C}^{-\beta}} \le c\|f\|_{C_T\mathcal{C}^\alpha}\|g\|_{C_T\mathcal{C}^{-\beta}}. \tag{2.2}$$

Below we recall some results on a class of PDEs with distributional drift in negative Besov spaces that will be used to set up the martingale problem for the singular SDE (1.1). All results are taken from [18]. In [18], as well as in the present work, the main assumption concerning the distribution-valued function b is the following.

Assumption A1 Let $0 < \beta < 1/2$ and $b \in C_T\mathcal{C}^{(-\beta)+}(\mathbb{R}^d)$. In particular $b \in C_T\mathcal{C}^{-\beta}(\mathbb{R}^d)$. Notice that b is a column vector.

We start by the formal definition of the operator $\mathcal{L}$.

Definition 2.2 (Definition 4.3, [18]) Let b satisfy Assumption A1. The operator $\mathcal{L}$ is defined as

$$\begin{array}{l} \mathcal{L} : \mathcal{D}_{\mathcal{L}}^0 \to \{\mathcal{S}'\text{-valued continuous functions}\} \\ \quad f \;\; \mapsto \mathcal{L} f := \dot{f} + \frac{1}{2}\Delta f + \nabla f\, b, \end{array}$$

where

$$\mathcal{D}_{\mathcal{L}}^0 := C_T D\mathcal{C}^{\beta} \cap C^1([0,T]; \mathcal{S}').$$

Here $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ and the function $\dot{f} : [0,T] \to \mathcal{S}'$ is the time-derivative of f. Note also that $\nabla f\, b := \nabla f \cdot b$ is well-defined using (2.1) and Assumption A1 and moreover it is continuous. The Laplacian Δ is intended in the sense of distributions.

Next we recall some results on certain PDEs, all driven by the operator $\mathcal{L}$. These results are all proved in the companion paper [18]. There are three equations of interest, all related but slightly different. The first PDE is

$$\begin{cases} \mathcal{L} v = g \\ v(T) = v_T. \end{cases} \tag{2.3}$$

We know from [18, Remark 4.8] that if $v_T \in \mathcal{C}^{(1+\beta)+}$ and $g \in C_T\mathcal{C}^{(-\beta)+}$ then there exists a unique (weak or mild) solution $v \in C_T\mathcal{C}^{(1+\beta)+}$. In [18, Lemma 4.17 and Remark 4.18] we prove a continuity result, namely that if the terminal condition v_T in (2.3) is replaced by a sequence (v_T^n) that converges to v_T in $\mathcal{C}^{(1+\beta)+}$, the terms b and g are replaced by two sequences (b^n) and (g^n), respectively, both converging in $C_T\mathcal{C}^{-\beta}$, then also the corresponding unique solutions (v^n) will converge to v in $C_T\mathcal{C}^{(1+\beta)+}$.

We can solve PDE (2.3) also under weaker conditions on v_T, in particular we allow functions with linear growth. The space that characterises this behaviour is denoted by $D\mathcal{C}^{\beta}$, which is the space of differentiable functions whose gradient belongs to $\mathcal{C}^{\beta}$. Notice that in [18] we introduce two concepts of solution, weak and mild, which are defined for functions in $C_T D\mathcal{C}^{\beta}$. We prove in [18, Proposition 4.5] that the notions of weak and mild solution of the PDE are equivalent. In [18, Remark 4.8] we show that if $v_T \in D\mathcal{C}^{\beta+}$ then there exists a unique solution $v \in C_T D\mathcal{C}^{\beta+}$. Continuity results for PDE (2.3) in the spaces $D\mathcal{C}^{\beta+}$ also hold, as we prove in [18, Remark 4.18 (i)], that is if $g^n \to g$ in $C_T\mathcal{C}^{-\beta}$, $b^n \to b$ in $C_T\mathcal{C}^{-\beta}$ and $v_T^n \to v_T$ in $D\mathcal{C}^{\beta+}$ then $v^n \to v$ in $C_T D\mathcal{C}^{\beta+}$. As a special case we show in [18, Corollary 4.10] that the function $\mathrm{id}_i(x) = x_i$ solves PDE (2.3) with $v(T) = x_i$ and $g = b_i$, that is $\mathcal{L}\mathrm{id}_i = b_i$.

Let $\lambda > 0$. The second PDE to consider is

$$\begin{cases} \mathcal{L}\phi_i = \lambda(\phi_i - \mathrm{id}_i) \\ \phi_i(T) = \mathrm{id}_i, \end{cases} \tag{2.4}$$

which has a unique (weak or mild) solution ϕ_i for $i = 1, \dots, d$ in the space $C_T D\mathcal{C}^{(1-\beta)-}$ (uniqueness holds in $C_T D\mathcal{C}^{\beta}$) by [18, Theorem 4.7 (i)]. In [18, Proposition 4.15] we show that $\phi_i \in \mathcal{D}_{\mathcal{L}}^0$ and $\dot{\phi}_i \in C_T\mathcal{C}^{(-\beta)-}$ for all $i = 1, \dots, d$. We denote by ϕ the column vector with components ϕ_i, $i = 1, \dots, d$. We show in [18, Proposition 4.16] that there exists $\lambda > 0$ large enough such that $\phi(t, \cdot)$ is invertible for all $t \in [0, T]$, and denoting such inverse with

$$\psi(t, \cdot) := \phi^{-1}(t, \cdot). \tag{2.5}$$

In the same proposition we also show that $\phi, \psi \in C^{0,1}$ and moreover that $\nabla\phi \in C_T\mathcal{C}^{(1-\beta)-}$ and $\nabla\psi(t, \cdot) \in \mathcal{C}^{(1-\beta)-}$ for all $t \in [0, T]$ and $\sup_{t\in[0,T]} \|\nabla\psi(t, \cdot)\|_\alpha$ for all $\alpha < 1-\beta$. From now on, let (b^n) be the sequence defined in [19, Proposition 2.4], so we know that $b^n \to b$ in $C_T\mathcal{C}^{-\beta}$, $b^n \in C_T\mathcal{C}^\gamma$ for all $\gamma > 0$ and b^n is bounded and Lipschitz. Here and in the rest of the paper $\lambda > 0$ is fixed and independent of n, chosen such that

$$\lambda = [C(\beta, \varepsilon) \max\{\sup_n \|b^n\|_{C_T\mathcal{C}^{-\beta+\varepsilon}}, \|b\|_{C_T\mathcal{C}^{-\beta+\varepsilon}}\}]^{\frac{1}{1-\theta}}, \tag{2.6}$$

according to [18, Lemma 4.19], where $\varepsilon > 0$ is such that $\theta := \frac{1+2\beta-\varepsilon}{2} < 1$ and $C(\beta, \varepsilon)$ is a constant only depending on β and ε. Notice with this choice of λ the corresponding inverse ψ^n of ϕ^n, see (2.5) is well-defined according to [18, Proposition 4.16 (ii)]. In [18, Lemma 4.19] we show that $\phi^n \to \phi$ and $\psi^n \to \psi$ uniformly on $[0, T] \times \mathbb{R}^d$ and $\|\nabla\phi^n\|_\infty + |\phi^n(0, 0)|$ is uniformly bounded in n.

Finally, in [18, Theorem 4.14] we show that the function ϕ is equivalently defined as $\phi = \mathrm{id} + u$, where $u = (u_1, \dots, u_d)$ and u_i is the unique solution of the third PDE, that is

$$\begin{cases} \mathcal{L}u_i = \lambda u_i - b_i \\ u_i(T) = 0 \end{cases} \tag{2.7}$$

in the space $C_T\mathcal{C}^{(2-\beta)-}$. For the latter PDE there are also continuity results proven in [18, Lemma 4.17], namely $u_i^n \to u_i$ in $C_T\mathcal{C}^{(2-\beta)-}$. Moreover, we have uniform convergence of $u^n \to u$, $\nabla u^n \to \nabla u$ by [18, Lemma 4.19]. With λ chosen as in (2.6) we have $\|\nabla u^n\|_\infty \le \frac{1}{2}$ by [18, Proposition 4.13 and bound (4.34)].

2.3 Probabilistic Notation

In the sequel we will consider generic measurable spaces $(\Omega, \mathcal{F})$. On them we will consider various probability measures denoted by $\mathbb{P}$. We will make use of the notation $(X, \mathbb{P})$ or $(Y, \mathbb{P})$, where X or Y will denote continuous stochastic processes indexed by $t \in [0, T]$ defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, without recalling it explicitly. The filtrations considered, if not explicitly mentioned, will be the canonical filtrations generated by X or Y (which will be the same in our applications).

Once the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is fixed, we will denote by $\mathscr{C}$ the linear space of continuous processes on $[0, T]$ with values in $\mathbb{R}^d$ endowed with the metric of uniform convergence in probability (u.c.p.).

The canonical space of continuous functions from $[0, T]$ with values in $\mathbb{R}^d$ will be denoted by $\mathcal{C}_T$, and it will be endowed with the sigma algebra of Borel sets $\mathcal{B}(\mathcal{C}_T)$. For $s \in [0, T]$ we will use the notation $\mathcal{C}_s$ for the space of continuous functions defined on

$[0, s]$. Thus, for a given couple $(X, \mathbb{P})$, the law of X under $\mathbb{P}$ will be a Borel probability measure on the measurable space $(\mathcal{C}_T, \mathcal{B}(\mathcal{C}_T))$.

3 A Zvonkin-type Transformation

In the study of SDEs with low-regularity coefficients, like (1.1), one successful idea is to apply a bijective transformation that changes the singular drift and produces a transformed SDE whose drift has no singular component and which can thus be solved with standard techniques. The idea goes back to Zvonkin [30], and in the present case a transformation that does the job is the unique solution ϕ of the PDE (2.4). The analysis that we do here can shed some light on what kind of transformations, aside from ϕ, of the martingale problem fulfilled by X will lead to different, but equivalent, transformed martingale problems fulfilled by a new process Y.

Let us start by introducing a class of function, denoted by $\mathcal{D}_{\mathcal{L}}$, that is the domain of the martingale problem

$$\begin{aligned} \mathcal{D}_{\mathcal{L}} := \{ f \in C_T\mathcal{C}^{(1+\beta)+} : \exists g \in C_T\bar{\mathcal{C}}_c^{0+} \text{ such that} \\ f \text{ is a weak solution of } \mathcal{L}f = g \text{ and } f(T) \in \bar{\mathcal{C}}_c^{(1+\beta)+} \}, \end{aligned} \tag{3.1}$$

where $\mathcal{L}$ has been defined in Definition 2.2.

Definition 3.1 We say that a couple $(X, \mathbb{P})$, where X is a continuous process indexed by $t \in [0, T]$ and $\mathbb{P}$ is a probability on some measurable space, is a *solution to the martingale problem with distributional drift b and initial condition μ* (for shortness, solution of MP with distributional drift b and i.c. μ) if and only if for every $f \in \mathcal{D}_{\mathcal{L}}$

$$f(t, X_t) - f(0, X_0) - \int_0^t (\mathcal{L}f)(s, X_s)\mathrm{d}s \tag{3.2}$$

is a local martingale under $\mathbb{P}$, and $X_0 \sim \mu$ under $\mathbb{P}$, where the domain $\mathcal{D}_{\mathcal{L}}$ is given by (3.1) and $\mathcal{L}$ has been defined in Definition 2.2.

We say that the martingale problem with distributional drift *b admits uniqueness* if for any two solutions $(X^1, \mathbb{P}^1)$ and $(X^2, \mathbb{P}^2)$ with $X_0^i \sim \mu$, $i = 1, 2$, then the law of X^1 under $\mathbb{P}^1$ is the same as the law of X^2 under $\mathbb{P}^2$.

Remark 3.2 Since $\bar{\mathcal{C}}_c^{(1+\beta)+} \subset \bar{\mathcal{C}}_c^{0+} \subset \mathcal{C}^{(-\beta)+}$, then there exists a unique weak solution $f \in C_T\mathcal{C}^{(1+\beta)+}$ for the PDE appearing in $\mathcal{D}_{\mathcal{L}}$, see Sect. 2.2. Moreover, by [18, Remark 4.4] we have $\mathcal{D}_{\mathcal{L}} \subset \mathcal{D}_{\mathcal{L}}^0$.

Proposition 3.3 *The domain $\mathcal{D}_{\mathcal{L}}$ defined in* (3.1) *equipped with its graph topology is separable.*

Proof By [18, Lemma 5.7 (i)] with $\gamma = 0$ we know that $\bar{\mathcal{C}}_c^{0+}$ is separable; hence, there exists a dense subset D_0 of $\bar{\mathcal{C}}_c^{0+}$, and by [18, Corollary 5.8] we know that $C_T\bar{\mathcal{C}}_c^{\beta+}$ is separable; thus, there exists a dense subset D_β of $C_T\bar{\mathcal{C}}_c^{\beta+}$. Let us denote by D the set of all $f_n \in C_T\mathcal{C}^{(1+\beta)+}$ such that $\mathcal{L}f_n = g_n$; $f_n(T) = f_n^T$ where $g_n \in D_0$ and $f_n^T \in D_\beta$.

Clearly, D is countable, because D_0 and D_β are countable and $D \subset \mathcal{D}_\mathcal{L}$. Moreover, by continuity results on the PDE (2.3), see Sect. 2.2, we have that if $f_n^T \to f(T)$ in $\mathcal{C}^{(1+\beta)+}$ and $g_n \to g$ in $C_T\mathcal{C}^{0+}$, then $f_n \to f$ in $C_T\mathcal{C}^{(1+\beta)+}$, which proves that the set D is dense in $\mathcal{D}_\mathcal{L}$. □

Next, we introduce the transformed SDE studied here, which is

$$Y_t = Y_0 + \lambda \int_0^t Y_s \mathrm{d}s - \lambda \int_0^t \psi(s, Y_s)\mathrm{d}s + \int_0^t \nabla\phi(s, \psi(s, Y_s))\mathrm{d}W_s, \tag{3.3}$$

where ϕ is the unique solution of (2.4) and ψ is its (space-)inverse given by (2.5) with $\lambda > 0$ chosen large enough (see Sect. 2.2). Notice that this SDE is formally obtained by applying the transformation ϕ to X as in Definition 3.1, that is, setting $Y_t = \phi(t, X_t)$ and using that ϕ is invertible with inverse ψ.

Denoting by Y the solution of (3.3), by Itô's formula for all $\tilde{f} \in C^{1,2}_{buc}([0, T] \times \mathbb{R}^d)$ we know that

$$\tilde{f}(t, Y_t) - \tilde{f}(0, Y_0) - \int_0^t (\tilde{\mathcal{L}}\tilde{f})(s, Y_s)\mathrm{d}s$$

is a martingale under $\mathbb{P}$. Here the operator $\tilde{\mathcal{L}}$ is the generator of Y, which is defined by

$$\tilde{\mathcal{L}}\tilde{f} := \partial_t \tilde{f} + \lambda\nabla\tilde{f}(\mathrm{id} - \psi) + \frac{1}{2}\mathrm{Tr}[(\nabla\phi \circ \psi)^\top \mathrm{Hess}\tilde{f}(\nabla\phi \circ \psi)]. \tag{3.4}$$

In particular, $(Y, \mathbb{P})$ verifies the classical Stroock–Varadhan martingale problem with respect to $\tilde{\mathcal{L}}$. We recall that this notion is equivalent to the one of weak solution for SDEs, see [20, Proposition 4.11 in Chapter 5]. To avoid confusion with the notion of weak solution for PDEs, in this paper we use the terminology *solution in law* instead of weak solution when referring to SDEs.

Remark 3.4 Note that the coefficients in $\tilde{\mathcal{L}}$ belong to $C^{0,\nu}$ for any $\nu < 1 - \beta$, see Appendix A for the definition of $C^{0,\nu}$. Indeed, $\tilde{f} \in C^{1,2}_{buc}$, ψ has linear growth since $|\nabla\psi|$ is uniformly bounded and the coefficient is $\nabla\phi \circ \psi$ belongs to $C^{0,\nu}$ for any $\nu < 1 - \beta$ because

$$\begin{aligned}
\|\nabla\phi(t, \psi(t, \cdot))\|_\nu \leq & \sup_x |\nabla\phi(t, \psi(t, x))| \\
& + \sup_{x_1, x_2,\, x_1 \neq x_2} \frac{|\nabla\phi(t, \psi(t, x_2)) - \nabla\phi(t, \psi(t, x_2))|}{|x_1 - x_2|^\nu} \\
\leq & \sup_{t\in[0,T]} \|\nabla\phi(t, \cdot)\|_\infty \\
& + \sup_{t\in[0,T]} \sup_{x_1, x_2,\, x_1 \neq x_2} \frac{|\nabla\phi(t, \psi(t, x_2)) - \nabla\phi(t, \psi(t, x_2))|}{|\psi(t, x_1) - \psi(t, x_2)|^\nu}
\end{aligned}$$

$$\frac{|\psi(t,x_1)-\psi(t,x_2)|^\nu}{|x_1-x_2|^\nu}$$
$$\le \sup_{t\in[0,T]} \|\nabla\phi(t,\cdot)\|_\infty + \|\nabla\phi\|_{C_T\mathcal{C}^\nu}\|\nabla\psi\|_\infty.$$

Here we have also used Remark A.1.

It will be useful later on to consider a domain for the operator $\tilde{\mathcal{L}}$ obtained as the image of $\mathcal{D}_\mathcal{L}$ through ϕ. Let us define

$$\tilde{\mathcal{D}}_{\tilde{\mathcal{L}}} := \{\tilde f = f\circ\psi \text{ for some } f\in\mathcal{D}_\mathcal{L} \text{ and } \psi \text{ defined in (2.5)}\}. \tag{3.5}$$

The choice of the SDE (3.3) and of the domain $\tilde{\mathcal{D}}_{\tilde{\mathcal{L}}}$ is natural since we use the transformed process $Y_t = \phi(t, X_t)$.

Lemma 3.5 *Let $g, h : \mathbb{R}^d \to \mathbb{R}^d$ with $h \in C^1$ with $\nabla h \in C^{\beta+}$ and $g \in \mathcal{C}^{(1+\beta)+}$. Then $g\circ h \in \mathcal{C}^{(1+\beta)+}$. If moreover $g_n \to g$ in $\mathcal{C}^{(1+\beta)+}$, then $g_n\circ h \to g\circ h$ in $\mathcal{C}^{(1+\beta)+}$.*

Proof To prove that $g\circ h \in \mathcal{C}^{(1+\beta)+}$ is equivalent to prove that $\bar f := g(h(\cdot))$ is bounded and that there exists $\alpha > \beta$ such that $\nabla\bar f \in \mathcal{C}^\alpha$, i.e. $\nabla\bar f$ is bounded and α-Hölder continuous. The first claim is obvious by boundedness of g. The gradient $\nabla\bar f(\cdot) = \nabla h(\cdot)\nabla g(h(\cdot))$ is bounded because it is the product of two bounded matrices since $\nabla h, \nabla g$ are bounded by assumption on g, h.

To show that $\nabla\bar f$ is α-Hölder continuous, it is enough to show that it is the product of two functions in $\mathcal{C}^\alpha$ (note that boundedness of the factors is crucially used). We have $\nabla h \in \mathcal{C}^\alpha$ for some $\alpha > \beta$ by assumption. On the other hand it is immediate to show that the term $\nabla g(h(\cdot))$ is in $\mathcal{C}^\alpha$, because it is bounded, and α-Hölder continuity is proved using that $\nabla g \in \mathcal{C}^\alpha$ and h is Lipschitz because by assumption ∇h is bounded.

To show convergence, let us denote $\bar f_n := g_n\circ h$. Since $\bar f_n(0) \to \bar f(0)$, it is enough to show the convergence of $\nabla\bar f_n$ in $\mathcal{C}^\alpha$. We use the same properties as above to get $\|\nabla\bar f_n - \nabla\bar f\|_\alpha \le \|\nabla h\|_\infty\|g_n - g\|_\alpha + \|\nabla g_n - \nabla g\|_\infty\|h\|_\alpha$, and the proof is complete. □

Lemma 3.6 *If $\tilde f \in C^{1,2}_{buc}$ and ϕ is the unique solution to PDE (2.4), then $\tilde f\circ\phi \in C_T\mathcal{C}^{(1+\beta)+}$.*

Proof Let us set $f := \tilde f\circ\phi$. We first prove that $f(t) \in \mathcal{C}^{(1+\beta)+}$ for all $t\in[0,T]$. This is a consequence of Lemma 3.5 with $g = \tilde f(t,\cdot)$ and $h = \psi(t,\cdot)$. The hypothesis on g is satisfied since $g \in C^{1,2}_{buc}$ and hence $g(t) \in \mathcal{C}^{(1+\gamma)+}$ for any $\gamma\in(0,1)$. The hypothesis on h is satisfied since $\nabla h \in \mathcal{C}^{(1-\beta)-}$ implies $\nabla h \in \mathcal{C}^{\beta+}$.

For the (uniform) time-continuity with values in $\mathcal{C}^{(1+\beta)+}$, since β is not an integer, we have to control

$$\|f(t) - f(s)\|_\infty + \|\nabla f(t) - \nabla f(s)\|_\alpha \tag{3.6}$$

for some $\alpha > \beta$ and for small $|t-s|$, where we recall $f(t) = \tilde f(t, \phi(t,\cdot))$, having used the equivalent norm [18, (2.3)]. The first term in (3.6) is obvious from the fact that $\tilde{f} \in C^{1,2}_{buc}$ and

$$\phi(t,x) - \phi(s,x) = u(t,x) - u(s,x), \quad \text{where } u \in C_T\mathcal{C}^{1+\alpha}, \tag{3.7}$$

see Sect. 2.2.

For the second term in (3.6), setting $H := \nabla\tilde{f} \circ \phi$, we can write $\nabla f = H\nabla\phi$. We note that $\nabla\phi \in C_T\mathcal{C}^\alpha$, see Sect. 2.2, and since $H \in C_T\mathcal{C}^\alpha$ (proved below) then the product is also in $C_T\mathcal{C}^\alpha$ and the proof is concluded.

It remains to show that $H \in C_T\mathcal{C}^\alpha$. For the sup part of the norm (see [18, (2.2)]), we notice that

$$\begin{aligned}
\|H(t) - H(s)\|_\infty &\le \|\nabla\tilde{f}(t,\phi(t,\cdot)) - \nabla\tilde{f}(t,\phi(s,\cdot))\|_\infty \\
&\quad + \|\nabla\tilde{f}(t,\phi(s,\cdot)) - \nabla\tilde{f}(s,\phi(s,\cdot))\|_\infty \\
&\le \|\mathrm{Hess}(\tilde{f})\|_\infty \|\phi(t,\cdot) - \phi(s,\cdot)\|_\infty \\
&\quad + \|\nabla\tilde{f}(t,\phi(s,\cdot)) - \nabla\tilde{f}(s,\phi(s,\cdot))\|_\infty,
\end{aligned} \tag{3.8}$$

and the first term is bounded as above using (3.7), while the second term is controlled because $\tilde{f} \in C^{1,2}_{buc}$.

We observe that $H \in C^{0,1}$ and $\nabla H = (\mathrm{Hess}(\tilde{f}) \circ \phi)\nabla\phi$. We will use below that ∇H is uniformly continuous, which we see by showing that each term of the product is bounded and uniformly continuous (*buc*). $\nabla\phi$ is *buc* because $\nabla\phi \in C_T\mathcal{C}^\alpha$. The term $\mathrm{Hess}(\tilde{f}) \circ \phi$ is similar to (3.8) but using that $\mathrm{Hess}(\tilde{f})$ is *buc* and (3.7).

Concerning the α-seminorm (see [18, (2.2)]), for x_1, x_2 such that $|x_1 - x_2| < 1$ we have

$$\begin{aligned}
&\frac{|H(t,x_1) - H(t,x_2) - (H(s,x_1) - H(s,x_2))|}{|x_1 - x_2|^\alpha} \\
&\le \int_0^1 |\nabla H(t, x_1 + a(x_2 - x_1)) - \nabla H(s, x_1 + a(x_2 - x_1))| \, \mathrm{d}a |x_2 - x_1|^{1-\alpha} \\
&\le \omega_{\nabla H}(|t - s|),
\end{aligned}$$

where $\omega_{\nabla H}(\cdot)$ denotes the continuity modulus of ∇H. This concludes the control of the second term in (3.6). □

Lemma 3.7 *Let $\tilde{f} \in C^{1,2}_{buc}$ and ϕ be the unique solution of PDE (2.4). Setting $f := \tilde{f} \circ \phi$ we have $f \in \mathcal{D}^0_{\mathcal{L}}$ and*

$$(\tilde{\mathcal{L}}\tilde{f}) \circ \phi = \mathcal{L}f$$

in $C_T\mathcal{C}^{0+}$, that is f is a solution of $\mathcal{L}f = g$ with $g := (\tilde{\mathcal{L}}\tilde{f}) \circ \phi \in C_T\mathcal{C}^{0+}$. Equivalently, we have $\tilde{\mathcal{L}}\tilde{f} = (\mathcal{L}f) \circ \psi$, where ψ is the space-inverse of ϕ defined in (2.5).

If moreover $\tilde{f}$ has compact support, then $f(T)$ and g also have compact support, in which case $f \in \mathcal{D}_{\mathcal{L}}$.

Proof We start by proving that $f \in \mathcal{D}^0_{\mathcal{L}}$ so that we can then calculate $\mathcal{L}f$. Notice that $f \in C_T\mathcal{C}^{(1+\beta)+}$ by Lemma 3.6. To show that $f \in C^1([0,T],\mathcal{S}')$, we compute the time-derivative $\dot{f}$. Recall that $\tilde{f} \in C^{1,2}_{buc}$ by assumption, and that $\phi : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ and $f, \tilde{f} : [0,T]\times\mathbb{R}^d \to \mathbb{R}$. We have

$$t \mapsto \dot{f}(t,\cdot) = \dot{\tilde{f}}(t,\phi(t,\cdot)) + \sum_{k=1}^{d} \partial_k \tilde{f}(t,\phi(t,\cdot))\dot{\phi}_k(t,\cdot), \tag{3.9}$$

where the dot $\dot{}$ denotes the time-derivative and $\partial_k := \frac{\partial}{\partial x_k}$. We show that the right-hand side of equation (3.9) is in $C_T\mathcal{S}'$. For the first term in (3.9) clearly we have the claim because $\dot{\tilde{f}} \circ \phi$ is uniformly continuous in t, x. The second term in (3.9) has products of the form $(\partial_k\tilde{f}\circ\phi)\dot{\phi}_k$ where $\dot{\phi}_k \in C_T\mathcal{C}^{(-\beta)-}$, see Sect. 2.2 and $\partial_k\tilde{f}\circ\phi \in C_T\mathcal{C}^{\beta+}$. Hence, the product is well-defined and continuous by (2.2). This shows that $f \in C^1([0,T];\mathcal{S}')$ and hence $f \in \mathcal{D}^0_{\mathcal{L}}$.

We now apply $\mathcal{L}$ to f so we need to calculate the spatial derivatives of f. The first space derivative of f with respect to x_i is

$$\partial_i f(t,\cdot) = \sum_{k=1}^{d} \partial_k\tilde{f}(t,\phi(t,\cdot))\partial_i\phi_k(t,\cdot), t \in [0,T]$$

and the second derivative is

$$\begin{aligned}\partial_{ii} f(t,\cdot) &= \sum_{k=1}^{d}\left[\sum_{l=1}^{d}(\partial_{lk}\tilde{f}(t,\phi(t,\cdot))\partial_i\phi_l(t,\cdot))\partial_i\phi_k(t,\cdot) + \partial_k\tilde{f}(t,\phi(t,\cdot))\partial_{ii}\phi_k(t,\cdot)\right]\\ &= \left((\nabla\phi)^T(\mathrm{Hess}(\tilde{f})\circ\phi)\nabla\phi\right)_{ii}(t,\cdot) + \sum_{k=1}^{d}\partial_k\tilde{f}(t,\phi(t,\cdot))\partial_{ii}\phi_k(t,\cdot), t \in [0,T].\end{aligned}$$

Note that $\partial_i f(t,\cdot)$ for all $t \in [0,T]$ is a well-defined object in $\mathcal{C}^{(-\beta)-}$ because it is actually a function in $\mathcal{C}^{\beta+}$ by Lemma 3.6. The second derivative $\partial_{ii} f(t,\cdot)$ is made of two terms: the first one is a bounded function, and the second one is well-defined in $\mathcal{C}^{(-\beta)-}$ again by means of the pointwise product (2.1), where for all $t \in [0,T]$ the distributional term $\partial_{ii}\phi_k(t,\cdot)$ is in $\mathcal{C}^{(-\beta)-}$ since $\partial_i\phi_k(t,\cdot) \in \mathcal{C}^{(1-\beta)-}$, see Sect. 2.2. Using these we calculate $\mathcal{L}\{$:

$$\begin{aligned}(\mathcal{L}f)(t,\cdot) =& \dot{\tilde{f}}(t,\phi(t,\cdot)) + \frac{1}{2}\sum_{i=1}^{d}\left((\nabla\phi)^T(\mathrm{Hess}(\tilde{f})\circ\phi)\nabla\phi\right)_{ii}(t,\cdot)\\ &+ \sum_{k=1}^{d}\partial_k\tilde{f}(t,\phi(t,\cdot))\left[\dot{\phi}_k(t,\cdot) + \frac{1}{2}\sum_{i=1}^{d}\partial_{ii}\phi_k(t,\cdot) + \partial_i\phi_k(t,\cdot)b_i(t,\cdot)\right], t \in [0,T],\end{aligned} \tag{3.10}$$

where the last term $\partial_k\tilde{f}(t,\phi(t,\cdot))\partial_i\phi_k(t,\cdot)b_i(t,\cdot)$ is well-defined in $\mathcal{C}^{(-\beta)+}$ by (2.1) used twice. Thus, equality (3.10) holds in the space $\mathcal{C}^{(-\beta)-}$. Now we observe that

$\mathcal{L}\phi_k = \lambda(\phi_k - \mathrm{id}_k)$ because ϕ_k is solution of PDE (2.4), see Sect. 2.2. Thus, the equality above becomes

$$
\begin{aligned}
(\mathcal{L}f)(t,\cdot) =& \dot{\tilde{f}}(t,\phi(t,\cdot)) + \frac{1}{2}\sum_{i=1}^{d}\Big((\nabla\phi)^T(\mathrm{Hess}(\tilde{f})\circ\phi)\nabla\phi\Big)_{ii}(t,\cdot)\\
&+\sum_{k=1}^{d}\partial_k\tilde{f}(t,\phi(t,\cdot))\lambda(\phi_k(t,\cdot)-\mathrm{id}_k)\\
=& \dot{\tilde{f}}(t,\phi(t,\cdot)) + \frac{1}{2}\mathrm{Tr}\Big((\nabla\phi)^T(\mathrm{Hess}(\tilde{f})\circ\phi)\nabla\phi\Big)(t,\cdot)\\
&+\lambda\nabla\tilde{f}(t,\phi(t,\cdot))(\phi(t,\cdot)-\mathrm{id}),\, t\in[0,T]. \qquad (3.11)
\end{aligned}
$$

On the other hand, by direct definition (3.4) of $\tilde{\mathcal{L}}$ applied to $\tilde{f} \in C^{1,2}_{buc}$ and then composed with ϕ and using $\psi(t,\phi(t,\cdot)) = \mathrm{id}$, one easily gets

$$
\begin{aligned}
(\tilde{\mathcal{L}}\tilde{f})(t,\phi(t,\cdot)) =& \dot{\tilde{f}}(t,\phi(t,\cdot)) + \frac{1}{2}\mathrm{Tr}\Big((\nabla\phi)^T(\mathrm{Hess}(\tilde{f})\circ\phi)\nabla\phi\Big)(t,\cdot)\\
&+\lambda\nabla\tilde{f}(t,\phi(t,\cdot))(\phi(t,\cdot)-\mathrm{id}),\, t\in[0,T]. \qquad (3.12)
\end{aligned}
$$

Now using (3.11) and (3.12) we get $t \mapsto (\mathcal{L}f) = (\tilde{\mathcal{L}}\tilde{f})(t,\phi(t,\cdot))$ in $C([0,T];\mathcal{S}')$. We observe that the right-hand side of (3.12) belongs to $C_T\mathcal{C}^{0+}$. Setting $g := (\tilde{\mathcal{L}}\tilde{f})\circ\phi$ we can conclude that $\mathcal{L}f = g \in C_T\mathcal{C}^{0+}$. Given that both sides are functions, we can compose them with ψ to get $\tilde{\mathcal{L}}\tilde{f} = (\mathcal{L}f)\circ\psi$.

Finally, we show that if $\tilde{f}$ has compact support, then $g = (\tilde{\mathcal{L}}\tilde{f})\circ\phi$ also has compact support. First notice that $\tilde{\mathcal{L}}\tilde{f}$ has compact support; thus, there exists $M > 0$ such that for all (t,x) with $|(t,\phi(t,x))| > M$ then $g(t,x) = 0$. To show that g has compact support it is enough to find $N > 0$ such that if $|(t,x)| > N$, then $|(t,\phi(t,x))| > M$. This is equivalent to showing that

$$
A := \{(t,x) : |(t,\phi(t,x))| \le M\} \subset \{(t,x) : |(t,x)| \le N\} =: B,
$$

for some N. To show the above inclusion, let $(t,x) \in A$. We write $(t,x) = (t,\psi(t,\phi(t,x)))$ and using that $\nabla\psi$ is uniformly bounded, see Sect. 2.2, we get

$$
\begin{aligned}
|(t,x)| &= |(t,\psi(t,\phi(t,x)) - \psi(t,0) + \psi(t,0))|\\
&\le C|(t,\phi(t,x))| + |(t,\psi(t,0))|\\
&\le CM + \sup_{t\in[0,T]}|(t,\psi(t,0))| =: N,
\end{aligned}
$$

which shows that $(t,x) \in B$. We conclude by noting that $f(T,\cdot)$ also has compact support, following the above computations but fixing the time $t = T$ and replacing $\tilde{\mathcal{L}}\tilde{f}$ with $\tilde{f}$. □

Lemma 3.8 *We have* $C^{1,2}_c \subset \tilde{\mathcal{D}}_{\tilde{\mathcal{L}}}$.

Proof By Definition of $\tilde{\mathcal{D}}_{\tilde{\mathcal{L}}}$ we have to show that if $\tilde{f} \in C_c^{1,2}$ then $f := \tilde{f} \circ \phi \in \mathcal{D}_{\mathcal{L}}$, where $\mathcal{D}_{\mathcal{L}}$ is given in (3.1). First, we note that by Lemma 3.6 we have $f \in C_T\mathcal{C}^{(1+\beta)+}$. Next we show that $\mathcal{L}\{ = g$ for some $g \in C_T\bar{\mathcal{C}}_c^{0+}$. We define $g := \tilde{\mathcal{L}}\tilde{f} \circ \phi$. By Lemma 3.7 we have $\mathcal{L}f = g$ and since $\tilde{f}$ has compact support, then $f \in \mathcal{D}_{\mathcal{L}}$ by Lemma 3.7 again. □

We can finally state the main result of this section, namely the equivalence between the original martingale problem and the Zvonkin-transformed martingale problem.

Theorem 3.9 *Let Assumption A1 hold.*

(i) If $(X, \mathbb{P})$ is a solution to MP with distributional drift b and i.c. μ then $(Y, \mathbb{P})$ is a solution in law to (3.3), where $Y_t := \phi(t, X_t)$ and $Y_0 \sim \nu$, where ν is the pushforward measure of μ given by $\nu := \mu(\psi(0, \cdot))$.

(ii) If $(Y, \mathbb{P})$ is a solution in law to (3.3) with $Y_0 \sim \nu$ then $(X, \mathbb{P})$ is a solution to MP with distributional drift b and i.c. μ, where $X_t := \psi(t, Y_t)$ and μ is the pushforward measure of ν given by $\mu := \nu(\phi(0, \cdot))$.

Proof *Item (i).* Let $(X, \mathbb{P})$ be a solution of MP. For any $\tilde{f} \in C_c^\infty$ we define $f := \tilde{f} \circ \phi$, where ϕ is the unique solution of PDE (2.4). By Lemma 3.7 $f \in \mathcal{D}_{\mathcal{L}}$. Setting $Y_t := \phi(t, X_t)$, by Lemma 3.7 we have

$$\tilde{f}(Y_t) - \tilde{f}(Y_0) - \int_0^t (\tilde{\mathcal{L}}\tilde{f})(s, Y_s)\mathrm{d}s = f(t, X_t) - f(0, X_0) - \int_0^t (\mathcal{L}f)(s, X_s)\mathrm{d}s,$$

which is a local martingale under $\mathbb{P}$ for all $\tilde{f} \in C_c^\infty$ by Definition 3.1 since $f \in \mathcal{D}_{\mathcal{L}}$. It follows that the couple $(Y, \mathbb{P})$ satisfies the Stroock–Varadhan martingale problem; therefore, $(Y, \mathbb{P})$ is a solution in law of SDE (3.3).

Item (ii). Let $(Y, \mathbb{P})$ be a solution in law of SDE (3.3). We define $X_t := \psi(t, Y_t)$, where ψ is the (space-)inverse of ϕ defined in (2.5). To show that $(X, \mathbb{P})$ is a solution to MP with distributional drift b, we need to show that for all $f \in \mathcal{D}_{\mathcal{L}}$ the quantity

$$f(t, X_t) - f(0, X_0) - \int_0^t (\mathcal{L}f)(s, X_s)\mathrm{d}s$$

is a local martingale under $\mathbb{P}$. Since $f \in \mathcal{D}_{\mathcal{L}}$ then there exists $g \in C_T\mathcal{C}^{0+}$ (so there exists $\nu \in (0, 1)$ with $g \in C_T\mathcal{C}^\nu$) such that $\mathcal{L}f = g$. We define $\tilde{g} := g \circ \psi$, $\tilde{f}_T := f(T, \psi(T, \cdot))$ and $\tilde{f}_T^n := \tilde{f}_T * \rho_n$, where $\rho_n = p_{\frac{1}{n}}$ with p_t the heat kernel. We see that $\tilde{g} \in \mathcal{C}^{0,\nu}$, see Appendix A for the explicit definition of the space. Indeed $\tilde{g}$ is in $C([0, T] \times \mathbb{R}^d)$ because g and ψ are, and it is easy to obtain the bound

$$\sup_{t\in[0,T]} \sup_{x\neq y} \frac{|\tilde{g}(t, x) - \tilde{g}(t, y)|}{|x - y|^\nu} \leq \sup_{t\in[0,T]} \|g(t)\|_{\mathcal{C}^\nu} \|\nabla\psi\|_\infty^\nu,$$

using the fact that $g \in C_T\mathcal{C}^\nu$ and $\psi \in C^{0,1}$ with gradient $\nabla\psi$ uniformly bounded, see Sect. 2.2. Moreover, $\tilde{f}_T^n \in C^{2+\nu}$ (for explicit definition of these spaces and its inclusion in other spaces, see Appendix A) and by Remark 3.4 the coefficients of $\tilde{\mathcal{L}}$ are in $C^{0,\nu}$. So by [21, Theorem 5.1.9] (which has been recalled in Theorem A.3 in the Appendix for ease of reading) we know that for each n there exists a function $\tilde{f}^n \in C^{1,2+\nu}([0,T]\times\mathbb{R}^d)$ (see Appendix A for the definition of this space and its inclusion in other spaces) which is the classical solution of

$$\begin{cases} \tilde{\mathcal{L}}\tilde{f}^n = \tilde{g} \\ \tilde{f}^n(T) = \tilde{f}^n_T. \end{cases} \tag{3.13}$$

Therefore, $\tilde{f}^n \in C^{1,2}$ and thus by Itô's formula

$$\tilde{f}^n(t, Y_t) - \tilde{f}^n(0, Y_0) - \int_0^t \tilde{g}(s, Y_s)\mathrm{d}s$$

is a local martingale under $\mathbb{P}$. Here we used that $(\tilde{\mathcal{L}}\tilde{f}_n)(s, Y_s) = \tilde{g}(s, Y_s)$ by construction. Setting $f^n := \tilde{f}^n \circ \phi$, we also have that

$$f^n(t, X_t) - f^n(0, X_0) - \int_0^t g(s, X_s)\mathrm{d}s \tag{3.14}$$

is a local martingale under $\mathbb{P}$. Using the definition of $\tilde{g}$, the fact that $\tilde{f}^n$ is a classical solution of PDE (3.13) and $\tilde{f}^n \in C^{1,2}_{buc}$ (see Remark A.2) by Lemma 3.7 we know that

$$g = \tilde{g} \circ \phi = \tilde{\mathcal{L}}\tilde{f}^n \circ \phi = \mathcal{L}f^n,$$

in $C_T\mathcal{C}^\nu$ and thus in particular f^n is a weak solution of

$$\begin{cases} \mathcal{L}f^n = g \\ f^n(T) = f^n_T, \end{cases} \tag{3.15}$$

where $f^n_T := \tilde{f}^n(T) \circ \phi(T, \cdot)$.

Now we claim that f^n is the unique mild solution to (3.15) in $C_T\mathcal{C}^{(1+\beta)+}$ and that $f^n \to f$ uniformly on compacts (these claims will be proven later). By this convergence and taking the limit of (3.14) where we replace $g = \mathcal{L}f$, we get that

$$f(t, X_t) - f(0, X_0) - \int_0^t (\mathcal{L}f)(s, X_s)\mathrm{d}s$$

is a local martingale under $\mathbb{P}$, thanks to the fact that the space of local martingales is closed under u.c.p. convergence.

It is left to prove that f^n is the unique mild solution to (3.15) in $C_T\mathcal{C}^{(1+\beta)+}$ and that $f^n \to f$ uniformly on compacts, which we do in three steps.

Step 1: we prove that f^n is the unique mild solution to (3.15) *in $C_T\mathcal{C}^{(1+\beta)+}$*. To do so, first we show that $f^n \in C_T\mathcal{C}^{(1+\beta)+}$, indeed $f^n := \tilde{f}^n \circ \phi$ with $\tilde{f}^n \in C^{1,2}_{buc}$ and ϕ solution of PDE (2.4), so by Lemma 3.6 we have $f^n \in C_T\mathcal{C}^{(1+\beta)+}$. In Sect. 2.2 it is recalled that weak and mild solutions are equivalent therefore f^n is the unique (mild) solution in $C_T\mathcal{C}^{(1+\beta)+}$.

Step 2: we prove that $f^n_T \to f_T := f(T)$ in $\mathcal{C}^{(1+\beta)+}$. Recall that $f^n_T = \tilde{f}^n_T \circ \phi(T,\cdot)$, so by Lemma 3.6 again we have $f^n_T \in \mathcal{C}^{(1+\beta)+}$. Moreover, $f_T = f(T) \in \mathcal{C}^{(1+\beta)+}$ because $f \in \mathcal{D}_\mathcal{L}$. Now we notice that $\tilde{f}_T \in \mathcal{C}^{(1+\beta)+}$ by Lemma 3.5 using the definition $\tilde{f}_T := f_T \circ \psi(T,\cdot)$, where $f_T \in \mathcal{C}^{(1+\beta)+}$ by definition of $\mathcal{D}_\mathcal{L}$ and $\psi(T,\cdot) \in C^1$ with $\nabla\psi(T,\cdot) \in \mathcal{C}^{(1-\beta)-}$ see Sect. 2.2. Since $\tilde{f}^n_T = \tilde{f}_T * \rho_n$ and the convolution with the mollifier ρ_n maintains the same regularity of $\tilde{f}_T$ by [18, Lemma 2.4], then $\tilde{f}^n_T \to \tilde{f}_T$ in $\mathcal{C}^{(1+\beta)+}$, see Sect. 2.2. Finally, again by Lemma 3.5 we have $\tilde{f}^n_T \circ \phi(T,\cdot) \to \tilde{f}_T \circ \phi(T,\cdot)$ in $\mathcal{C}^{(1+\beta)+}$ as wanted.

Step 3: we prove that $f^n \to f$ uniformly, in particular uniformly on compacts. From Step 1 we have that f^n is the unique solution of (3.15) in $C_T C^{(1+\beta)+}$. Moreover, we recall that f is the unique mild solution in the same space of $\mathcal{L}f = g$ with terminal condition the value of the function itself, $f_T = f(T)$. We can now apply continuity results on the PDE (3.15), see Sect. 2.2, to conclude that $f^n \to f$ in $C_T\mathcal{C}^{(1+\beta)+}$. This clearly implies that $f^n \to f$ uniformly, as wanted. □

Remark 3.10 It is possible to define an equivalent MP by a transformation different than the one used in Theorem 3.9. Indeed, it is enough to consider a generic transformation $\phi \in C_T D\mathcal{C}^{\beta+}$ which is space-invertible with inverse ψ, and under which one has the equivalence between $(X,\mathbb{P})$ solving the MP with respect to $\mathcal{L}$ and $(\phi(X),\mathbb{P})$ solving the MP with respect to $\tilde{\mathcal{L}}$, where $\tilde{\mathcal{L}}\tilde{f} := \mathcal{L}f \circ \psi$. The issue going further would be to interpret $\tilde{\mathcal{L}}\tilde{f} = \tilde{g}$ as a PDE, which would need to be considered in the mild sense and will presumably require some regularity of ϕ. Well-posedness of such an equation would be based on Schauder-type estimates for the time-dependent semigroup generated by the diffusive component of the operator $\tilde{\mathcal{L}}$, which are far from being straightforward.

From now on, let (b^n) be the sequence defined in [19, Proposition 2.4], so we know that $b^n \to b$ in $C_T\mathcal{C}^{-\beta}$, $b^n \in C_T\mathcal{C}^\gamma$ for all $\gamma \in \mathbb{R}$ and b^n is bounded and Lipschitz. Recall that $\lambda > 0$ has been fixed and independent of n, chosen such that (2.6) holds. To conclude the section, we prove a continuity result for the transformed problem for Y that will be useful when we will prove analogous continuity results for the original problem for X. Let us denote by Y^n the strong solution of

$$Y^n_t = \phi(0, X_0) + \lambda \int_0^t Y^n_s \mathrm{d}s - \lambda \int_0^t \psi^n(s, Y^n_s)\mathrm{d}s + \int_0^t \nabla\phi^n(s, \psi^n(s, Y^n_s))\mathrm{d}W_s, \tag{3.16}$$

which is the counterpart of (3.3) when one replaces b with b^n.

Remark 3.11 We notice that the drift and the diffusion coefficient of (3.16) are uniformly bounded in n. Indeed the drift is given by $\lambda(y - \psi^n(s, y)) = \lambda u^n(s, \psi^n(s, y))$ and the diffusion coefficient is $\nabla\phi^n(s, \psi^n(s, y)) = \nabla u^n(s, \psi^n(s, y)) + I_d$. Thanks [18, Lemma 4.9], for every fixed $\alpha \in (\beta, 1-\beta)$ we have

$$\|u^n\|_{C_T\mathcal{C}^{\alpha+1}} \le R_\lambda(\|b^n\|_{C_T\mathcal{C}^{-\beta}})\|b^n\|_{C_T\mathcal{C}^{-\beta}} \le R_\lambda(\sup_n \|b^n\|_{C_T\mathcal{C}^{-\beta}}) \sup_n \|b^n\|_{C_T\mathcal{C}^{-\beta}},$$

where R_λ is an increasing function. Thus, u_n and ∇u_n are uniformly bounded in n.

Lemma 3.12 *Let Y^n be the solution of SDE (3.16). Then the sequence of laws of (Y^n) is tight.*

Proof According to [20, Theorem 4.10 in Chapter 2] we need to prove that

$$\lim_{\eta\to\infty}\sup_{n\geq 1}\mathbb{P}(|Y_0^n|>\eta)=0 \tag{3.17}$$

and that for every $\varepsilon>0$

$$\lim_{\delta\to 0}\sup_{n\geq 1}\mathbb{P}\Big(\sup_{\substack{s,t\in[0,T]\\|s-t|\leq\delta}}|Y_t^n-Y_s^n|>\varepsilon\Big)=0. \tag{3.18}$$

We know that $Y_0^n=\phi^n(0,X_0)$ and $X_0\sim\mu$. By continuity results on the PDE (2.4), see Sect. 2.2, we have that $\phi^n\to\phi$ uniformly and that

$$a:=\sup_{n\geq 1}\|\nabla\phi^n\|_\infty<\infty \quad \text{and} \quad b:=\sup_{n\geq 1}|\phi^n(0,0)|<\infty.$$

So the first condition (3.17) for tightness gives

$$\begin{aligned}\mathbb{P}(|Y^n(0)|>\eta)&=\mathbb{P}(|\phi^n(0,X_0)|>\eta)\\&\leq\mathbb{P}(|\phi^n(0,0)|+\|\nabla\phi^n\|_\infty|X_0|>\eta)\\&\leq\mathbb{P}(a+b|X_0|>\eta).\end{aligned}$$

Noticing that $a+b|X_0|$ is a finite random variable (independent of n) then we have (3.17).

Concerning the second bound (3.18) for tightness, we first observe that the classical Kolmogorov criterion

$$\mathbb{E}[|Y_t^n-Y_s^n|^4]\leq C|t-s|^2 \tag{3.19}$$

holds for some positive constant C independent of n. The proof of this bound works exactly as the proof in [12, Step 3 of Proposition 29]: indeed, the process Y^n therein has the same form as Y^n given by (3.16). By Remark 3.11 we have that the drift and diffusion coefficients are uniformly bounded in n, so that [12, Step 3 of Proposition 29] allows to show (3.19).

Now we apply Garsia–Rodemich–Rumsey Lemma (see e.g. [3, Sect. 3]) and we know that for every $0<m<1$ there exists a constant C' and a random variable Γ_n such that

$$|Y_t^n-Y_s^n|^4\leq C'|t-s|^m\Gamma_n$$

with

$$\mathbb{E}(\Gamma_n) \leq c\, C \frac{1}{1-m} T^{2-m}, \tag{3.20}$$

where c is a universal constant. Consequently, for every $\varepsilon > 0$ and for every $n \geq 1$

$$\begin{aligned}
\mathbb{P}\Big(\sup_{\substack{s,t\in[0,T]\\ |s-t|\leq\delta}} |Y^n_t - Y^n_s| > \varepsilon \Big) &= \mathbb{P}\Big(\varepsilon < \sup_{\substack{s,t\in[0,T]\\ |s-t|\leq\delta}} |Y^n_t - Y^n_s| \leq C'^{\frac{1}{4}} \delta^{\frac{m}{4}} \Gamma_n^{\frac{1}{4}} \Big) \\
&\leq \mathbb{P}\Big(\varepsilon \leq C'^{\frac{1}{4}} \delta^{\frac{m}{4}} \Gamma_n^{\frac{1}{4}} \Big) \\
&\leq \mathbb{P}\Big(\Gamma_n \geq \frac{\varepsilon^4}{C'\delta^m} \Big) \\
&\leq \frac{C'\delta^m}{\varepsilon^4} \mathbb{E}(\Gamma_n),
\end{aligned}$$

by Chebyshev inequality. So, using (3.20) we have that $\sup_{n\geq 1} \mathbb{P}\Big(\sup_{\substack{s,t\in[0,T]\\ |s-t|\leq\delta}} |Y^n_t - Y^n_s| > \varepsilon \Big) \to 0$ as $\delta \to 0$ and (3.18) is established. □

Remark 3.13 When $Y_0 = y$ is a deterministic initial condition, we know that (3.3) admits existence and uniqueness in law by [28, Theorem 10.2.2], because the drift and diffusion coefficient are bounded by Remark 3.11 and the diffusion coefficients is continuous since $\nabla\phi$ and ψ are continuous and it is uniformly non-degenerate since $\|\nabla u\|_\infty \leq \frac{1}{2}$, see Sect. 2.2.

4 The Martingale Problem for *X*

In this section we solve the martingale problem for the process X, which formally satisfies an SDE of the form

$$X_t = X_0 + \int_0^t b(s, X_s)\mathrm{d}s + W_t,$$

where W is a d-dimensional Brownian motion, the drift b is an element of $C_T\mathcal{C}^{(-\beta)+}$ that satisfies Assumption A1 and the initial condition X_0 is a given random variable. To do so, we first solve the problem for a deterministic initial condition and then we use this to extend the result to any initial condition. We also derive some properties about said solution, such as its link to the Fokker–Planck equation and continuity properties.

We start with the case when the drift b is a function, by comparing the notion of solution to the singular MP with the notion of solution in law of SDEs, and with the Stroock–Varadhan Martingale Problem, see [28, Sect. 6.0]. We recall that $(X, \mathbb{P})$ is a solution to the Stroock–Varadhan Martingale Problem with respect to $\mathcal{L}$ if for every $f \in C^\infty_c$

$$f(X_t) - f(X_0) - \int_0^t (\frac{1}{2}\Delta f(X_s) + \nabla f(X_s) b(s, X_s)) \mathrm{d}s \tag{4.1}$$

is a local martingale.

Lemma 4.1 *Let $b \in C_T \mathcal{C}^{0+}$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be some probability space. Let $X_0 \sim \mu$. Then the following are equivalent.*

(i) The couple $(X, \mathbb{P})$ is solution to the MP with distributional drift b.
(ii) The couple $(X, \mathbb{P})$ is solution to the Stroock–Varadhan Martingale Problem with respect to $\mathcal{L}$.
(iii) There exists a Brownian motion W such that the process X under $\mathbb{P}$ is a solution of $\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \mathrm{d}W_t$.

Proof (ii) $\iff$ (iii). This follows from the Stroock–Varadhan classical theory, see [28, Chapter 8]. We sketch the proof for completeness. If the Stroock–Varadhan Martingale Problem is fulfilled, i.e. if (ii) holds, then in fact (4.1) also holds for $f \in C^2$. Choosing $f(x) = x^i$ and $f(x) = x^i x^j, 1 \le i, j \le d$, one can show that $M_t = X_t - X_0 - \int_0^t b(s, X_s)\mathrm{d}s$ is a local martingale with covariation matrix $([X^i, X^j])_{i,j}$ being the identity. The process M is then a standard d-dimensional Brownian motion by Lévy's characterisation theorem. Vice versa if X fulfils the SDE (iii) then (ii) follows by Itô's formula.

(i) $\Longrightarrow$ (ii). For this it is enough to show that for every $f \in C_c^\infty$ (4.1) holds. This is true since $C_c^\infty \subset \mathcal{D}_{\mathcal{L}}$ in this case.

(iii) $\Longrightarrow$ (i). We will make use of the spaces $C^{0,\nu}([0, T] \times \mathbb{R}^d)$ and $C^{1,2+\nu}$ for $\nu \in (0, 1)$, which have been defined in Appendix A. Since $b \in C_T \mathcal{C}^{0+}$, by [18, Remark 4.12] we know that the unique solution $u \in C_T \mathcal{C}^{(1+\beta)+}$ of PDE (2.7) is also the classical solution as given in Theorem A.3, hence $u \in C^{1,2}$. We set $\phi = \mathrm{id} + u$, which thus belongs to $C^{1,2}$ so by Itô's formula applied to $Y = \phi(t, X_t)$ where X is a solution to $\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \mathrm{d}W_t$ we get that Y solves (3.3) with initial condition $Y_0 \sim \nu := \mu(\psi(0, \cdot))$, where ψ is the inverse of ϕ. Thus, Theorem 3.9 implies that $(X, \mathbb{P})$ is a solution to the MP with (distributional) drift b and i.c. μ, as wanted. □

Next we show the link between the law of the solution to the MP and the Fokker–Planck equation, in particular we show that the law of the solution to the martingale problem with distributional drift satisfies a Fokker–Planck equation.

Theorem 4.2 *Let Assumption A1 hold. Let $(X, \mathbb{P})$ be a solution to the martingale problem with distributional drift b and initial condition μ with density v_0. Let $v(t, \cdot)$ be the law density of X_t and let us assume that $v \in C_T \mathcal{C}^{\beta+}$. Then v is a weak solution of the Fokker–Planck equation, that is for every $\varphi \in \mathcal{S}$ we have*

$$\langle \varphi, v(t) \rangle = \langle \varphi, v_0 \rangle + \int_0^t \langle \frac{1}{2}\Delta\varphi, v(s) \rangle \mathrm{d}s + \int_0^t \langle \nabla\varphi, v(s) b(s) \rangle \mathrm{d}s, \tag{4.2}$$

for all $t \in [0, T]$.

Notice that the product $v(s)b(s)$ appearing in the last integral is well-defined using pointwise products (2.1). We remark that the solution v is the unique solution of (4.2) by [19, Theorem 3.7 and Proposition 3.2].

Proof It is enough to show the claim for all $\varphi \in C_c^\infty$. Indeed C_c^∞ is dense in $\mathcal{S}$. Since $\varphi \in C_c^\infty \subset \mathcal{D}_\mathcal{L}^0$, then we can apply the operator $\mathcal{L}$ defined in Definition 2.2 to φ, and we define $\mathcal{L}\varphi =: g$. Clearly φ is a weak solution of the PDE $\mathcal{L}\varphi = g$ with terminal condition φ. Moreover, the function φ is time independent by construction. Using the definition of $\mathcal{L}$ we get for all $s \in [0, T]$ that

$$(\mathcal{L}\varphi)(s) = \tfrac{1}{2}\Delta\varphi + \nabla\varphi\, b(s) \tag{4.3}$$

in $\mathcal{C}^{-\beta}$ (having used the regularity of φ and the pointwise product (2.1)). In fact since $t \mapsto b(t, \cdot) \in \mathcal{C}^{-\beta}$ is a continuous function of time by (2.2) we have that $\mathcal{L}\varphi \in C_T\mathcal{C}^{-\beta}$.

We now construct a sequence $(g^n) \in C_T\mathcal{C}^{0+}$ that converges to g in $C_T\mathcal{C}^{-\beta}$ and that is compactly supported. Let (b^n) be the sequence defined before (2.6), in particular it converges to b in $C_T\mathcal{C}^{-\beta}$ and let us define $g^n := \frac{1}{2}\Delta\varphi + \nabla\varphi\, b^n$. Then clearly $g^n \in C_T\mathcal{C}^{0+}$ (in fact it is more regular) and

$$\|g - g^n\|_{C_T\mathcal{C}^{-\beta}} = \|\nabla\varphi\,(b - b^n)\|_{C_T\mathcal{C}^{-\beta}} \le \|\nabla\varphi\|_{C_T\mathcal{C}^{\beta+}}\|b - b^n\|_{C_T\mathcal{C}^{-\beta}},$$

and the right-hand side goes to 0 as $n \to \infty$. Moreover, denoting by K the compact support of φ, we have that also g^n is supported on K.

Let us denote by u^n the mild solution of $\mathcal{L}u^n = g^n$, $u^n(T) = \varphi$, which exists and is unique in $C_T\mathcal{C}^{(1+\beta)+}$, see Sect. 2.2. Such function belongs to $\mathcal{D}_\mathcal{L}$ by definition of the domain $\mathcal{D}_\mathcal{L}$, see (3.1). Since $u^n \in \mathcal{D}_\mathcal{L}$ and $(X, \mathbb{P})$ is a solution to the martingale problem with distributional drift b and initial condition μ with density v_0, then we know that

$$u^n(t, X_t) - u^n(0, X_0) - \int_0^t \mathcal{L}u^n(s, X_s)\mathrm{d}s$$

is a local martingale under $\mathbb{P}$, but also a true martingale since u^n and $\mathcal{L}u^n$ are bounded. We denoted by $v(t, \cdot)$ the law density of X_t; thus, taking the expectation under $\mathbb{P}$ we have

$$\int_{\mathbb{R}^d} u^n(t, x)v(t, x)\mathrm{d}x - \int_{\mathbb{R}^d} u^n(0, x)v_0(x)\mathrm{d}x - \int_0^t \int_{\mathbb{R}^d} (\mathcal{L}\sqcap^n)(s, x)v(s, x)\mathrm{d}x\mathrm{d}s = 0. \tag{4.4}$$

We now consider a smooth function $\chi_K \in C_c^\infty$ such that $\chi_K = 1$ on K. Since g^n is compactly supported on K and $\mathcal{L}\sqcap^n = g^n$, we can rewrite the double integral in (4.4) as

$$\begin{aligned}\int_0^t \int_{\mathbb{R}^d} (\mathcal{L}\sqcap^n)(s, x)v(s, x)\mathrm{d}x\mathrm{d}s &= \int_0^t \int_{\mathbb{R}^d} (\mathcal{L}\sqcap^n)(s, x)v(s, x)\chi_K(x)\mathrm{d}x\mathrm{d}s \\ &= \int_0^t \langle(\mathcal{L}\sqcap^n)(s)v(s), \chi_K\rangle\mathrm{d}s,\end{aligned}$$

where the dual pairing is in $\mathcal{S}, \mathcal{S}'$. By continuity properties of the PDE $\mathcal{L}u^n = g^n$ with terminal condition $u^n(T) = \varphi$ (see Sect. 2.2) we know that since $g^n \to g$ in $C_T\mathcal{C}^{-\beta}$ then $u^n \to \varphi$ in $C_T\mathcal{C}^{(1+\beta)+}$, since φ is the unique solution of $\mathcal{L}\varphi = g$ with terminal condition $u(T) = \varphi$. Thus, taking the limit as $n \to \infty$ of the above dual pairing, we get

$$\begin{aligned}\lim_{n\to\infty}\int_0^t \langle(\mathcal{L}\sqcap^n)(s)v(s), \chi_K\rangle \mathrm{d}s &= \int_0^t \langle(\mathcal{L}\varphi)(s)v(s), \chi_K\rangle \mathrm{d}s \\ &= \int_0^t \langle\frac{1}{2}\Delta\varphi\, v(s), \chi_K\rangle + \langle\nabla\varphi\, b(s)v(s), \chi_K\rangle \mathrm{d}s \\ &= \int_0^t \langle\frac{1}{2}\Delta\varphi, v(s)\rangle \mathrm{d}s + \int_0^t \langle\nabla\varphi\, b(s)v(s), \chi_K\rangle \mathrm{d}s.\end{aligned} \tag{4.5}$$

Now we prove that the latter dual pairing in (4.5) can be rewritten as

$$\langle\nabla\varphi\, b(s)v(s), \chi_K\rangle = \langle\nabla\varphi, b(s)v(s)\rangle, \tag{4.6}$$

for all $s \in [0, T]$. Indeed, the LHS of (4.6) is well-defined because $\chi_K \in C_c^\infty$ and for every $s \in [0, T]$ the distribution $\nabla\varphi\, b(s)v(s)$ is actually an element of $\mathcal{C}^{-\beta}$ because of the pointwise product (2.1) and of the regularity $v(s) \in \mathcal{C}^{\beta+}$ and $b(s) \in \mathcal{C}^{-\beta}$. The RHS of (4.6) is also well-defined, but now the test function is $\nabla\varphi \in C_c^\infty$ and the distribution is $b(s)v(s)$. To show that (4.6) holds we observe that by the continuity of the product (2.2) we have $b^n(s)v(s) \to b(s)v(s)$ in $\mathcal{C}^{-\beta}$ (in fact uniformly in $s \in [0, T]$) and thus we can write

$$\begin{aligned}\langle\nabla\varphi\, b(s)v(s), \chi_K\rangle &= \lim_{n\to\infty}\langle\nabla\varphi\, b^n(s)v(s), \chi_K\rangle \\ &= \lim_{n\to\infty}\int_{\mathbb{R}^d} \nabla\varphi(x)b^n(s,x)v(s,x)\chi_K(x)\mathrm{d}x \\ &= \lim_{n\to\infty}\int_{\mathbb{R}^d} \nabla\varphi(x)b^n(s,x)v(s,x)\mathrm{d}x \\ &= \lim_{n\to\infty}\langle\nabla\varphi, b^n(s)v(s)\rangle \\ &= \langle\nabla\varphi, b(s)v(s)\rangle,\end{aligned}$$

for all $s \in [0, T]$, which proves (4.6).

To conclude it is enough to take the limit as $n \to \infty$ in (4.4) and use (4.5) and (4.6) to get (4.2). □

The following is a continuity result for the martingale problem. Recall that (b^n) is the sequence defined before (2.6) in Sect. 2.2, so we know that $b^n \to b$ in $C_T\mathcal{C}^{-\beta}$, $b^n \in C_T\mathcal{C}^\gamma$ for all $\gamma \in \mathbb{R}$ and b^n is bounded and Lipschitz. We denote by X^n the (strong) solution to the SDE

$$X_t^n = X_0 + \int_0^t b^n(s, X_s^n)\mathrm{d}s + W_t, \tag{4.7}$$

where $X_0 \sim \mu$.

Theorem 4.3 *Let Assumptions A1 hold. Let (b^n) be a sequence in $C_T\mathcal{C}^{(-\beta)+}$ converging to b in $C_T\mathcal{C}^{-\beta}$. Let $(X, \mathbb{P})$ (respectively $(X^n, \mathbb{P}^n)$) be a solution to the MP with distributional drift b (respectively b^n) and initial condition μ. Then the sequence $(X^n, \mathbb{P}^n)$ converges in law to $(X, \mathbb{P})$. In particular, if $b^n \in C_T\mathcal{C}^{0+}$ and X^n is a strong solution of*

$$X_t^n = X_0 + \int_0^t b^n(s, X_s^n)\mathrm{d}s + W_t,$$

then X^n converges to $(X, \mathbb{P})$ in law.

Proof The proof is identical to that of [12, Proposition 29]. In particular Step 4 therein deals with the convergence in law of Y^n, which is the solution of SDE (3.16), and Step 5 with the convergence in law of X^n. Notice that the drift b therein lives in a different space than ours (Bessel potential spaces instead of Hölder-Besov spaces), and the initial condition in [12] is deterministic, but the setting is otherwise the same. The only tools used in Step 4 and 5 are the tightness of the sequence of laws of Y^n, which we proved in Lemma 3.12, and the uniform convergence of $u^n \to u$, $\nabla u^n \to \nabla u$ and $\psi^n \to \psi$, see Sect. 2.2. Finally, setting $X_t := \psi(t, Y_t)$ for $t \in [0, T]$, then $(X, \mathbb{P})$ is the unique solution to the martingale problem with distributional drift b and initial condition μ by Theorem 3.9, because $(Y, \mathbb{P})$ is the unique solution to (3.3) with initial condition $Y_0 \sim \nu$ where ν is the pushforward measure of μ through ϕ.

It remains to prove the last claim of the theorem, which follows because X^n is also a solution to the MP with distributional drift b^n by Lemma 4.1, so the first part of the theorem can be applied. □

The first existence and uniqueness result is for the solution to the MP with distributional drift b and deterministic initial condition $X_0 = x$. We will extend the result to any random variable in Theorem 4.5.

Proposition 4.4 *The martingale problem with distributional drift b and i.c. δ_x, for $x \in \mathbb{R}^d$, admits existence and uniqueness according to Definition 3.1.*

Proof Let $(X, \mathbb{P})$ be a solution to the MP. Setting $Y_t = \phi(t, X_t)$ and $Y_0 = y := \phi(0, x)$, by Item (i) of Theorem 3.9 we have that $(Y, \mathbb{P})$ is a solution in law to (3.3). By Remark 3.13 the solution $(Y, \mathbb{P})$ is unique; hence, the law of X under $\mathbb{P}$ is uniquely determined.

Existence follows from the fact that equation (3.3) with $Y_0 = y$ has a solution in law, say $(Y, \mathbb{P})$, again by Remark 3.13. Then setting $X_t := \psi(t, Y_t)$ by Item (ii) of Theorem 3.9, we know that $(X, \mathbb{P})$ is a solution in law to MP with distributional drift b and i.c. δ_x. □

Next we extend the existence and uniqueness result of Proposition 4.4 to the general case when the initial condition X_0 is a random variable rather than a deterministic point.

Theorem 4.5 *Let Assumption A1 hold and let μ be a probability measure on $\mathbb{R}^d$. Then there exists a unique solution $(X, \mathbb{P})$ to the martingale problem with distributional drift b and initial condition μ.*

Proof *Existence.* The idea is to use a *superposition* argument in order to glue together the solutions of MP with a deterministic initial condition x, for all possible initial conditions x. This is implemented using the process $Y_t = \phi(t, X_t)$.

We have the measure μ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ which is the law of the initial condition X_0 and we define a new measure ν on the same space given by $\nu(B) = \mu(\psi(0, B))$ for any $B \in \mathcal{B}(\mathbb{R}^d)$. Notice that ν is the pushforward of μ through the function ϕ, where $\psi = \phi^{-1}$ has been defined in (2.5); thus, ν plays the role of the initial condition for the process $\phi(t, X_t)$. Let Y be the canonical process and $\mathbb{P}^y$ be a law of the canonical process on $\mathcal{C}_T$ such that $(Y, \mathbb{P}^y)$ is the unique weak solution to (3.3) with $Y_0 = y$. Then it is known by [28, Theorem 7.1.6] that $(y, C) \mapsto \mathbb{P}^y(C)$ is a random kernel for $y \in \mathbb{R}^d$ and $C \in \mathcal{B}(\mathcal{C}_T)$; hence, the probability $\mathbb{P}$ given by

$$\mathbb{P}(C) := \int \mathbb{P}^y(C)\nu(\mathrm{d}y) \tag{4.8}$$

is well-defined. Setting $X_t := \psi(t, Y_t)$, our candidate solution to the MP with distributional drift b and initial condition μ is $(X, \mathbb{P})$. First, we observe that for any $C \in \mathcal{B}(\mathcal{C}_T)$ of the form $C = \{\omega : \omega_0 \in B\}$ with some $B \in \mathcal{B}(\mathbb{R}^d)$, we have

$$\mathbb{P}^y(C) = \mathbb{P}^y(\omega \in C) = \mathbb{P}^y(Y_0 \in B) = \mathbb{1}_B(y), \tag{4.9}$$

having used that $\mathbb{P}^y$-a.s. the canonical process Y is such that $Y_0 = y$. This will allow us to show that the initial condition X_0 has law μ. Indeed, for any $A \in \mathcal{B}(\mathbb{R}^d)$ we set $B = \phi(0, A)$ and we calculate

$$\mathbb{P}(X_0 \in A) = \mathbb{P}(\psi(0, Y_0) \in A) = \mathbb{P}(Y_0 \in \phi(0, A)) = \mathbb{P}(Y_0 \in B). \tag{4.10}$$

Now using the definition (4.8) of $\mathbb{P}$ and setting $C = \{Y_0 \in B\}$ we have $\mathbb{P}(Y_0 \in B) = \mathbb{P}(C) = \int \mathbb{P}^y(C)\nu(\mathrm{d}y)$ and by (4.9) we have

$$\mathbb{P}(Y_0 \in B) = \int_B \nu(\mathrm{d}y) = \nu(B) = \nu(\phi(0, A)). \tag{4.11}$$

Finally, using the definition of ν and the fact that ψ is the inverse of ϕ we have $\mathbb{P}(X_0 \in A) = \mathbb{P}(C) = \mu(A)$ as wanted.

Next we show that for every $f \in \mathcal{D}_\mathcal{L}$ the process

$$M_u^f(X) := f(u, X_u) - f(0, X_0) - \int_0^u (\mathcal{L}f)(r, X_r)\mathrm{d}r, \tag{4.12}$$

is a martingale under $\mathbb{P}$, that is for every $f \in \mathcal{D}_\mathcal{L}$ and F_s bounded and continuous functional on $\mathcal{C}_s$ (see Sect. 2.3) we have

$$\mathbb{E}[M_t^f(X)F_s(X)] = \mathbb{E}[M_s^f(X)F_s(X)],$$

for all $0 \le s \le t \le T$. Indeed, we notice that under $\mathbb{P}^y$ we have $Y_0 \sim \delta_y$; hence, $X_0 \sim \delta_{\psi(0,y)} =: \delta_x$. Moreover, $(Y, \mathbb{P}^y)$ is a solution of (3.3) with i.c. $Y_0 = y$; hence, by Theorem 3.9 part (ii) we have that $(X_\cdot := \psi(\cdot, Y_\cdot), \mathbb{P}^y)$ is a solution to the MP with distributional drift b and i.c. $X_0 \sim \delta_x$; thus, by the definition of $\mathbb{P}$ given in (4.8) we get

$$\mathbb{E}[(M_t^f(X) - M_s^f(X))F_s(X)] = \int \mathbb{E}^y[(M_t^f(X) - M_s^f(X))F_s(X)]\nu(\mathrm{d}y) = 0,$$

where we denoted by $\mathbb{E}^y$ the expectation under $\mathbb{P}^y$.

Uniqueness. Here the idea is to use *disintegration* in order to reduce the MP to MPs with deterministic initial condition. We proceed by stating and proving two preliminary facts.

Fact 1 Let E^1 be a dense countable set in $C_c(\mathbb{R})$, E^2 be a dense countable set in $\mathcal{D}_\mathcal{L}$ and $E^{\mathcal{C}_s}$ be a countable set of bounded continuous functionals such that for every bounded continuous functional $F_s \in \mathcal{C}_s$ there exists a sequence $(F_s^n) \subset E^{\mathcal{C}_s}$ such that $F_s^n \to F_s$ in a pointwise uniformly bounded way, see (4.17). A couple $(X, \mathbb{P})$ is a solution to the MP with distributional drift b and initial condition X_0 if and only if

$$\mathbb{E}[M_t^f(X)F_s(X)g(X_0)] = \mathbb{E}[M_s^f(X)F_s(X)g(X_0)], \tag{4.13}$$

for every $f \in E^2$, $F_s \in E^{\mathcal{C}_s}$, $g \in E^1$ and $s < t$ with $s, t \in \mathbb{Q} \cap [0, T]$, where $M_u^f(X)$ is given by (4.12).

This fact can be seen as follows. First, we notice that, since M^f are bounded processes, if M^f is a local martingale, then it is also a true martingale, and hence, the MP with distributional drift is equivalent to (4.13) for all $f \in \mathcal{D}_\mathcal{L}$, $F_s \in \mathcal{C}_s$ and $g \in C_c$ and $s < t$ with $s, t \in [0, T]$.

Next, one can show that this is equivalent when choosing $s < t, s, t \in \mathbb{Q} \cap [0, T]$. Indeed for any bounded and continuous functional F_s on $\mathcal{C}_s$, for a sequence of rational times $s_n \downarrow s$ with $s < s_n < t$, we can associate a sequence of bounded and continuous functionals F_{s_n} on $\mathcal{C}_{s_n}$ by setting $F_{s_n}(\eta) := F_s(\eta|_{[0,s]})$ for $\eta \in \mathcal{C}_{s_n}$.

This allows to replace the condition $s \in \mathbb{Q} \cap [0, T]$ with $s \in [0, T]$. In order to replace $t \in \mathbb{Q} \cap [0, T]$ with $t \in [0, T]$, we choose a rational sequence $t_n \in (s, T]$ such that $t_n \to t$ and use the fact that the local martingale $t \mapsto M_t^f$ is continuous and Lebesgue dominated convergence theorem.

Finally, we use again Lebesgue dominated convergence theorem to see the validity of (4.13) for all $f \in \mathcal{D}_\mathcal{L}$, $F_s \in \mathcal{C}_s$ and $g \in C_c$ and $s < t$ with $s, t \in \mathbb{Q} \cap [0, T]$.

We remark that E^1 exists because C_c is separable by [18, Lemma 5.7 (ii)], E^2 exists because $\mathcal{D}_\mathcal{L}$ is separable by Proposition 3.3 and $E^{\mathcal{C}_s}$ exists by Lemma 4.6, whose statement and proof has been postponed at the end of this section.

Fact 2 Let $(X, \mathbb{P})$ be a solution to the MP with distributional drift b and i.c. μ. There exists a random kernel $\mathbb{P}^x$ such that $\mathbb{P} = \int \mathbb{P}^x \mathrm{d}\mu(x)$, where for μ-almost all $x \in \mathbb{R}^d$, $\mathbb{P}^x$ lives on $\{\omega \in \Omega : X_0(\omega) = x\}$ and for any bounded and continuous functional $G : C[0, T] \to \mathbb{R}$ we have

$$\mathbb{E}(G(X)) = \int_{\mathbb{R}^d} \mathbb{E}^x(G(X))\mathrm{d}\mu(x), \tag{4.14}$$

where $\mathbb{E}$ and $\mathbb{E}^x$ stand for the expectation under $\mathbb{P}$ and $\mathbb{P}^x$ respectively. This follows from the disintegration theorem in [8, Chapter III, nos. 70–72].

We now proceed with the proof of uniqueness. Let $(X^1, \mathbb{P}_1)$ and $(X^2, \mathbb{P}_2)$ be two solutions to the MP with distributional drift b and initial condition $X_0 \sim \mu$. Without loss of generality we can suppose that $X^1 = X^2 = X$ is the canonical process on $\Omega = \mathcal{C}_T$. Since $(X^i, \mathbb{P}_i)$, $i = 1, 2$ is a solution of the MP, then by Fact 1 we have

$$\mathbb{E}_i[(M_t^f(X) - M_s^f(X))F_s(X)g(X_0)] = 0,$$

for all $0 \leq s \leq t \leq T$, $s, t \in \mathbb{Q}$, $g \in E^1$, $f \in E^2$ and $F_s \in E^{\mathcal{C}_s}$ and $i = 1, 2$. We now apply Fact 2 to both $\mathbb{P}_1$ and $\mathbb{P}_2$, and in particular (4.14) with $G(\eta) = (M_t^f(\eta) - M_t^f(\eta))F_s(\eta)g(\eta_0)$ to rewrite the above equality as

$$\int_{\mathbb{R}^d} \mathbb{E}_i^x[(M_t^f(X) - M_s^f(X))F_s(X)g(X_0)]\mathrm{d}\mu(x) = 0, \tag{4.15}$$

for all $0 \leq s \leq t \leq T$, $s, t \in \mathbb{Q}$, $g \in E^1$, $f \in E^2$ and $F_s \in E^{\mathcal{C}_s}$ and $i = 1, 2$. Now we recall that for μ-almost all x, we have $X_0(\omega) = x$, $\mathbb{P}_i^x$-a.s.; thus, equation (4.15) becomes

$$\int_{\mathbb{R}^d} g(x)\mathbb{E}_i^x[(M_t^f(X) - M_s^f(X))F_s(X)]\mathrm{d}\mu(x) = 0,$$

for every $0 \leq s \leq t \leq T$, $s, t \in \mathbb{Q}$, $g \in E^1$, $f \in E^2$ and $F_s \in E^{\mathcal{C}_s}$ and $i = 1, 2$. Since g is arbitrarily chosen in a dense set of $C_c(\mathbb{R})$, then we have

$$\mathbb{E}_i^x[(M_t^f(X) - M_s^f(X))F_s(X)] = 0 \quad \mu\text{-a.e.}, \tag{4.16}$$

for every $0 \leq s \leq t \leq T$, $s, t \in \mathbb{Q}$, $f \in E^2$ and $F_s \in E^{\mathcal{C}_s}$ and $i = 1, 2$. Note that (4.16) is true because the sets $\mathbb{Q} \cap [0, T]$, E^2 and $E^{\mathcal{C}_s}$ are countable. By Fact 1 this means that the couple $(X, \mathbb{P}_i^x)$ is a solution to the MP with distributional drift b and initial condition δ_x, for $i = 1, 2$ for μ-almost all x. By Proposition 4.4 we have uniqueness of the MP with deterministic initial condition δ_x; hence, for μ-almost all x we have $\mathbb{P}_1^x = \mathbb{P}_2^x$. Thus, recalling the disintegration $\mathbb{P}_i = \int \mathbb{P}_i^x \mathrm{d}\mu(x)$ for $i = 1, 2$ from Fact 2, we conclude $\mathbb{P}_1 = \mathbb{P}_2$ as wanted. □

We conclude the section with the proof of a technical result used in Fact 1 in the proof of Theorem 4.5.

Lemma 4.6 *There exists a countable family D of bounded and continuous functionals from $C([0, T]; \mathbb{R}^d)$ to $\mathbb{R}$ such that any bounded and continuous functional*

$F : C([0,T];\mathbb{R}^d) \to \mathbb{R}$ can be approximated by a sequence $(F_n) \subset D$ in a pointwise uniformly bounded way, that is

$$\begin{aligned} &F_n \to F \text{ pointwise} \\ &\sup_n \sup_{\eta \in C([0,T];\mathbb{R}^d)} |F_n(\eta)| < \infty. \end{aligned} \tag{4.17}$$

Proof We set $T = 1$ without loss of generality. Let $\eta \in C([0,1];\mathbb{R}^d)$. By [18, Lemma 5.5] we know that the function $t \mapsto F(\eta(t))$ can be approximated by $F_n(\eta(\cdot)) := F(B_n(\eta,\cdot))$, where (B_n) are the $\mathbb{R}^d$-valued Bernstein polynomials defined for any function $\eta \in C([0,1];\mathbb{R}^d)$ by

$$B_n(\eta,t) := \sum_{j=0}^{n} \eta(\frac{j}{n}) t^j (1-t)^{n-j} \binom{n}{j}.$$

Notice that the convergence is uniform in t. Now for fixed n and $y_0, y_1, \ldots, y_n \in \mathbb{R}^d$ we consider the function f on $\mathbb{R}^{(n+1)d}$ defined by

$$f(y_0, y_1, \ldots, y_n) := F\left(\sum_{j=0}^{n} y_j (\cdot)^j (1-\cdot)^{n-j} \binom{n}{j}\right),$$

so that $F_n(\eta) = f(\eta(\frac{0}{n}), \eta(\frac{1}{n}), \ldots, \eta(\frac{n}{n}))$. Notice that $\sup_{\eta \in C[0,T]} |F_n(\eta)| \le \|F\|_\infty$. We have thus reduced the problem to approximating any continuous bounded function $f : \mathbb{R}^{(n+1)d} \to \mathbb{R}$. We further reduce the problem to continuous functions on $[-M,M]^{(n+1)d}$ by restriction, for some $M > 0$. Indeed, a function $f : [-M,M]^{(n+1)d} \to \mathbb{R}$ can be naturally extended to a bounded continuous function $\hat{f}$ on $\mathbb{R}^{(n+1)d}$ by setting for $x \in \mathbb{R}^{(n+1)d}$

$$\hat{f}(x) = f(x_1 \vee (-M) \wedge M, \ldots, x_{(n+1)d} \vee (-M) \wedge M).$$

One can see that $C([-M,M]^{(n+1)d})$ is separable by Stone–Weierstrass theorem. We denote by $D_{n,\mathrm{fin}}$ the dense set in the set of bounded and continuous functions from $\mathbb{R}^{(n+1)d} \to \mathbb{R}$.

The proof is concluded by setting $D := \cup_{n \in \mathbb{N}} D_n$, where

$$D_n := \Big\{ F : C([0,1];\mathbb{R}^d) \to \mathbb{R} : F(\eta) = f(\eta(\tfrac{0}{n}), \eta(\tfrac{1}{n}), \ldots, \eta(\tfrac{n}{n})), \\ \eta \in C([0,1];\mathbb{R}^d) \text{ for some } f \in D_{n,\mathrm{fin}} \Big\},$$

which is a countable set of bounded functions. Then for any bounded and continuous functional $F : C([0,1];\mathbb{R}^d) \to \mathbb{R}$ we construct the sequence (F_n) that converges to F pointwisely by choosing the appropriate element $F_n \in D_n$. Since the convergence in $C([-M,M]^{(n+1)d})$ is uniform, we also have $\sup_n \sup_{\eta \in C([0,1];\mathbb{R}^d)} |F_n(\eta)| < \infty$. □

Remark 4.7 One could also define the domain $\mathcal{D}_{\mathcal{L}}$ of the martingale problem as a subset of the smaller space $C_T\mathcal{C}^{(2-\beta)-}$ instead of the larger space $C_T\mathcal{C}^{(1+\beta)+}$. On the other hand, one could enlarge the domain by choosing functions with linear growth, namely in $C_T D\mathcal{C}^{\beta+}$. In both cases the analysis of the resulting MP is similar and should lead to an equivalent problem to the one studied in the present paper. We leave these details to the interested reader.

5 The Solution of the MP as Weak Dirichlet Process

In this section we focus on the weak Dirichlet decomposition property of the solution of the MP, which will be useful in Sect. 6 to characterise it as a solution of a generalised SDE. We notice that a solution to the martingale problem with distributional drift b is not a semimartingale in general. Indeed already in the fully studied case of dimension $d = 1$, see [14, Corollary 5.11], one sees that the solution is a semimartingale if and only if b is a Radon measure. We can, however, discuss and investigate other properties of this process, which turns out to be a weak Dirichlet process, and we identify the martingale component of the weak Dirichlet decomposition.

We start with the definition of weak Dirichlet process, that can be found in [15], see also [10, 11].

Definition 5.1 Let X be a continuous stochastic process on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{F}^X$ denote its canonical filtration.

- A process $\mathscr{A}$ is said to be an $\mathcal{F}^X$*-martingale orthogonal process* if $[N, \mathscr{A}] = 0$ for every $\mathcal{F}^X$-continuous local martingale N.
- The process X is said $\mathcal{F}^X$-*weak Dirichlet* if it is the sum of an $\mathcal{F}^X$-local martingale M and an $\mathcal{F}^X$-martingale orthogonal process $\mathscr{A}$.
  When $\mathscr{A}_0 = 0$ a.s., we call $X = M + \mathscr{A}$ the standard decomposition.

Remark 5.2
- The two equalities in the statement of Definition 5.1, that is $[N, \mathscr{A}] = 0$ and $X = M + \mathscr{A}$, are meant up to indistinguishability with respect to $\mathbb{P}$.
- The standard decomposition of a $\mathcal{F}^X$-weak Dirichlet process is unique.

In the remainder of the section, we let $(X, \mathbb{P})$ be the solution to the martingale problem with distributional drift b and initial condition μ, with $\mathbb{P}$ being a probability measure on some measurable space $(\Omega, \mathcal{F})$ that will be fixed throughout. We will make use of the space of processes $\mathscr{C}$, introduced in Sect. 2.3. Let Assumption A1 hold.

Proposition 5.3 *Let $f \in C^{0,1}([0, T] \times \mathbb{R}^d)$. Then $f(t, X_t)$ is an $\mathcal{F}^X$-weak Dirichlet process. In particular, X is an $\mathcal{F}^X$-weak Dirichlet process.*

Proof We recall that by Theorem 3.9 $X = \psi(t, Y_t)$ where $\psi \in C^{0,1}$ and (Y_t) is an $\mathcal{F}^X$-semimartingale. Then $f(t, X_t) = (f \circ \psi)(t, Y_t)$ is a $C^{0,1}$ function of a semimartingale; hence, it is a weak Dirichlet process by [15, Corollary 3.11]. □

From now on we denote by $f(t, X_t) = M^f + \mathscr{A}_t^f$ the standard decomposition of the weak Dirichlet process $f(t, X_t)$ for $f \in C^{0,1}$.

In what follows, we compute the covariation process between two martingale parts M^f and M^h, for two functions $f, h \in C^{0,1}$. To do so we first need some preparatory lemmata dealing with functions in some subspace of $\mathcal{D}_{\mathcal{L}}$. We denote by $\mathcal{D}^s_{\mathcal{L}}$ the space given by

$$\mathcal{D}^s_{\mathcal{L}} := \{f \text{ such that } \exists \tilde{f} \in C^{1,2}_c \text{ and } f = \tilde{f} \circ \phi\}, \tag{5.1}$$

which is obviously and algebra. Moreover, it is a linear subspace of $\mathcal{D}_{\mathcal{L}}$ by Lemma 3.8.

Proposition 5.4 *For $f, h \in \mathcal{D}^s_{\mathcal{L}}$ we have*

$$\mathcal{L}(fh) = (\mathcal{L}f)h + (\mathcal{L}h)f + \nabla f \nabla h. \tag{5.2}$$

Proof Let $f, h \in \mathcal{D}^s_{\mathcal{L}}$ and let us compute the time derivative of the product fh. We have

$$\partial_t(fh) = h\partial_t f + f\partial_t h, \tag{5.3}$$

which makes sense as we see below. Indeed, $h\partial_t f$ is well-defined because $h \in C_T\mathcal{C}^{(1+\beta)+}$ and $\partial_t f = \mathcal{L}f - \frac{1}{2}\Delta f - \nabla f b$ is an element of $C_T\mathcal{C}^{(\beta-1)+}$. The latter holds because $\mathcal{L}f \in C_T\mathcal{C}^{0+}$, $\frac{1}{2}\Delta f \in C_T\mathcal{C}^{(\beta-1)+}$ and $\nabla f b \in C_T\mathcal{C}^{-\beta}$, with $(\beta - 1) \leq -\beta$. Similarly for $f\partial_t h$.

We also calculate the Laplacian of fh

$$\tfrac{1}{2}\Delta(fh) = \frac{1}{2}(h\Delta f + 2\nabla f \nabla h + f\Delta h), \tag{5.4}$$

where we recall that $\nabla f \nabla h := \nabla f \cdot \nabla h$, and we calculate the transport term

$$b\nabla(fh) = b\nabla f\, h + b\nabla h\, f, \tag{5.5}$$

which are well-defined by similar arguments. Collecting (5.3), (5.4) and (5.5) then equality (5.2) follows. □

Lemma 5.5 *Let $f, h \in \mathcal{D}^s_{\mathcal{L}}$. Then*

$$[M^f, M^h]_t = \int_0^t (\nabla f)(s, X_s)(\nabla h)(s, X_s)\mathrm{d}s. \tag{5.6}$$

Proof By Proposition 5.4, $fh \in \mathcal{D}^s_{\mathcal{L}} \subset \mathcal{D}_{\mathcal{L}}$, so using the martingale problem, Proposition 5.3 (and considerations below) together with the uniqueness of the standard weak Dirichlet decomposition we have

$$(fh)(t, X_t) = M^{fh} + \int_0^t \mathcal{L}(fh)(s, X_s)\mathrm{d}s, \tag{5.7}$$

having incorporated the initial condition $(fh)(0, X_0)$ in the martingale part M^{fh} so that $\mathscr{A}_t^{fh} = \int_0^t \mathcal{L}(fh)(s, X_s)\mathrm{d}s$; hence, $\mathscr{A}_0^{fh} = 0$ as required. It holds also

$$f(t, X_t) = M_t^f + \int_0^t \mathcal{L}f(s, X_s)\mathrm{d}s \tag{5.8}$$

$$h(t, X_t) = M_t^h + \int_0^t \mathcal{L}h(s, X_s)\mathrm{d}s. \tag{5.9}$$

Integrating by parts $(fh)(t, X_t)$ and using (5.8) and (5.9), we have

$$\begin{aligned}(fh)(t, X_t) &= \int_0^t f(s, X_s)\mathrm{d}h(s, X_s) + \int_0^t h(s, X_s)\mathrm{d}f(s, X_s) + [f(\cdot, X), h(\cdot, X)]_t \\ &= \mathscr{M}_t + \int_0^t f(s, X_s)(\mathcal{L}\langle\rangle)(s, X_s)\mathrm{d}s + \int_0^t h(s, X_s)(\mathcal{L}\{\})(s, X_s)\mathrm{d}s + [M^f, M^h]_t,\end{aligned} \tag{5.10}$$

where $(\mathscr{M}_t)$ is some local martingale. Equations (5.7) and (5.10) give two decompositions of the semimartingale $(fh)(t, X_t)$. By uniqueness of the decomposition and taking into account Proposition 5.4, the conclusion (5.6) follows. □

Remark 5.6 We notice that both sides of (5.6) are well-defined also for $f, h \in C^{0,1}$.

Lemma 5.7 *$\mathcal{D}_{\mathcal{L}}^s$ is dense in $C^{0,1}([0, T] \times \mathbb{R}^d)$.*

Proof Let $\chi : \mathbb{R} \to \mathbb{R}_+$ be a smooth function such that

$$\chi(x) = \begin{cases} 0 & x \geq 0 \\ 1 & x \leq -1 \\ \in (0, 1) & x \in (-1, 0). \end{cases}$$

We set $\chi_n : \mathbb{R}^d \to \mathbb{R}$ as $\chi_n(x) := \chi(|x| - (n + 1))$. In particular

$$\chi_n(x) = \begin{cases} 0 & |x| \geq n + 1 \\ 1 & |x| \leq n \\ \in (0, 1) & \text{otherwise.} \end{cases}$$

Let $f \in C^{0,1}$. Let us define $\tilde{f} := f \circ \psi \in C^{0,1}$ and $\tilde{f}_n := \tilde{f}\chi_n$. Since $\tilde{f}_n \to \tilde{f}$ in $C^{0,1}$ also $f_n := \tilde{f}_n \circ \phi \to f$ in $C^{0,1}$; hence, we reduce to the case where $\tilde{f} = f \circ \psi$ has compact support.

We set

$$\tilde{f}_m(t, x) := m \int_t^{t+\frac{1}{m}} (f \star \rho_m)(s, x)\mathrm{d}s,$$

where ρ_m is a sequence of mollifiers with compact support and $\star$ denotes the space-convolution. Then $\tilde{f}_m \in C_c^{1,\infty}([0, T] \times \mathbb{R}^d)$ and $\tilde{f}_m \to \tilde{f}$ in $C^{0,1}$; hence, $f_m := \tilde{f}_m \circ \phi \to f$ in $C^{0,1}$. □

Theorem 5.8 *Let $f, h \in C^{0,1}$. Then*

$$[M^f, M^h]_t = \int_0^t (\nabla f)(s, X_s)(\nabla h)(s, X_s)\mathrm{d}s. \tag{5.11}$$

Proof First, we notice that (5.11) holds for every $f, h \in \mathcal{D}^s_{\mathcal{L}}$ by Lemma 5.5. Each side of (5.11) is well-defined for $f, h \in C^{0,1}$, by Remark 5.6. Moreover, by Lemma 5.7 $\mathcal{D}^s_{\mathcal{L}} \subset C^{0,1}$ is a dense subspace.

Next we show that, for fixed $h \in \mathcal{D}^s_{\mathcal{L}}$, the map $f \mapsto [M^f, M^h]$ is continuous and linear from $C^{0,1}$ to $\mathcal{C}$. For this we make use of Banach–Steinhaus theorem for F-spaces, see e.g. [9, Theorem 2.1]. Indeed, the space $C^{0,1}$ is clearly an F-space, and so is the linear space of continuous processes $\mathscr{C}$ equipped with the u.c.p. topology. Let $[M^f, M^h]^\varepsilon$ denote the ε-regularisation of the bracket $[M^f, M^h]$, see [26, Definition 4.2] or [24, Sect. 1] for a precise definition. Let $h \in \mathcal{D}^s_{\mathcal{L}}$ be fixed. The operator $T^\varepsilon : f \mapsto [M^f, M^h]^\varepsilon$ is linear and continuous from $C^{0,1}$ to $\mathscr{C}$. Finally, $[M^f, M^h]$ is well-defined as a u.c.p.-limit of $[M^f, M^h]^\varepsilon$, see [24, Proposition 1.1]. Thus, by Banach–Steinhaus the map $f \mapsto [M^f, M^h]$ is continuous from $C^{0,1}$. Since both members of (5.11) are continuous and linear, then (5.11) extends to all $f \in C^{0,1}$ and $h \in \mathcal{D}^s_{\mathcal{L}}$.

Finally, let $f \in C^{0,1}$ be fixed. By the same reasoning as above we extend (5.11) to $h \in C^{0,1}$. □

Corollary 5.9 *The map $f \mapsto \mathscr{A}^f$ is continuous (and linear) from $C^{0,1}$ to $\mathscr{C}$.*

Proof Since $f_n \to 0$ in $C^{0,1}$, then $f_n(\cdot, X) \to 0$ u.c.p. By Theorem 5.8 $[M^{f_n}] \to 0$, and taking into account [20, Chapter 1, Problem 5.25] we have that $M^{f_n} \to 0$ u.c.p. Using the decomposition $f_n(\cdot, X) = M^{f_n} + \mathscr{A}^{f_n}$, we have $\mathscr{A}^{f_n} \to 0$ u.c.p. and the proof is concluded. □

Remark 5.10 Let $\mathrm{id}_i(x) = x_i$. Then $\mathrm{id}_i \in C^{0,1}$. Setting $M^{\mathrm{id}} = (M^{\mathrm{id}_1}, \dots, M^{\mathrm{id}_d})^\top$ then by Theorem 5.8, we have

$$[M^{\mathrm{id}_i}, M^{\mathrm{id}_j}]_t = \delta_{ij} t.$$

Hence, by Lévy characterisation theorem this implies that $M^{\mathrm{id}} - X_0$ is a standard d-dimensional Brownian motion. We denote this Brownian motion by W^X.

Proposition 5.11 *For $f \in C^{0,1}([0, T] \times \mathbb{R}^d)$, we have*

$$M^f_t = f(0, X_0) + \int_0^t \nabla f(s, X_s) \cdot \mathrm{d}M^{id}_s.$$

Proof Recall that we write

$$f(t, X_t) = M^f_t + \mathscr{A}^f_t, \tag{5.12}$$

where the right-hand side is the standard (unique) decomposition of the left-hand side, as an $\mathcal{F}^X$-weak Dirichlet process. In particular $\mathscr{A}^f$ is an $\mathcal{F}^X$-orthogonal process with $\mathscr{A}_0^f = 0$ and M^f is the martingale component. We define $\tilde{\mathscr{A}}^f$ so that

$$f(t, X_t) = f(0, X_0) + \int_0^t \nabla f(s, X_s) \cdot \mathrm{d}M^{\mathrm{id}} + \tilde{\mathscr{A}}_t^f.$$

We will prove later that

$$[\tilde{\mathscr{A}}^f, N] = 0 \text{ for all continuous local} \mathcal{F}^X\text{-martingales } N. \tag{5.13}$$

From (5.13) we have that $\tilde{\mathscr{A}}^f$ is an $\mathcal{F}^X$-martingale orthogonal process with $\tilde{\mathscr{A}}_0^f = f(0, X_0) - f(0, X_0) = 0$; thus, by uniqueness of the decomposition of weak Dirichlet processes it must be $\tilde{\mathscr{A}}^f = \mathscr{A}^f$ and therefore

$$M_t^f = f(0, X_0) + \int_0^t \nabla f(s, X_s) \cdot \mathrm{d}M_s^{\mathrm{id}},$$

as wanted. It remains to prove (5.13). By definition of $\tilde{\mathscr{A}}^f$ and (5.12) we have

$$\begin{aligned} [\tilde{\mathscr{A}}^f, N]_t &= [f(\cdot, X_\cdot), N]_t - [\int_0^\cdot \nabla f(s, X_s) \cdot \mathrm{d}M^{\mathrm{id}}, N]_s \\ &= [M^f, N]_t - \int_0^t \nabla f(s, X_s) \cdot \mathrm{d}[M^{\mathrm{id}}, N]_s, \end{aligned} \tag{5.14}$$

having used the weak Dirichlet decomposition $f(\cdot, X) = M^f + \mathscr{A}^f$, where $\mathscr{A}^f$ is an $\mathcal{F}^X$-martingale orthogonal process. Regarding N, now we observe that by Kunita–Watanabe decomposition there is an $\mathcal{F}^X$-progressively measurable process ξ and an orthogonal local martingale O such that

$$N_t = N_0 + \int_0^t \xi_s \cdot dM_s^{\mathrm{id}} + O_t.$$

Thus, the covariation with M^{id} gives

$$[M^{\mathrm{id}}, N]_t = [M^{\mathrm{id}}, \int_0^\cdot \xi_s \cdot \mathrm{d}M_s^{\mathrm{id}}]_t = \int_0^t \xi_s \mathrm{d}s,$$

since $[M^{\mathrm{id}_i}, M^{\mathrm{id}_j}]_t = \delta_{i,j} t$ by Remark 5.10. We calculate $[M^f, N]_t$ using Theorem 5.8 to get

$$[M^f, N]_t = [M^f, \int_0^\cdot \xi_s \cdot \mathrm{d}M_s^{\mathrm{id}}]_t = \int_0^t \xi_s \cdot \mathrm{d}[M^f, M^{\mathrm{id}}]_s = \int_0^t \xi_s \cdot \nabla f(s, X_s) \mathrm{d}s.$$

Plugging these two covariations into (5.14), we get

$$[\tilde{\mathscr{A}}^f, N]_t = \int_0^t \xi_s \cdot \nabla f(s, X_s)\mathrm{d}s - \int_0^t \nabla f(s, X_s) \cdot \xi_s \mathrm{d}s = 0,$$

which is (5.13) as wanted. □

We conclude this section with some final remarks.

Remark 5.12 (i) We recall that $\mathcal{D}^s_{\mathcal{L}} \subset \mathcal{D}_{\mathcal{L}} \subset C^{0,1}$. Thus, for $f \in \mathcal{D}^s_{\mathcal{L}}$ by uniqueness of the weak Dirichlet decomposition and by the martingale problem we have $\mathscr{A}^f_t = \int_0^t (\mathcal{L}f)(s, X_s)\mathrm{d}s$. Therefore, we have that $f \mapsto \mathscr{A}^f$ is the continuous linear extension of $f \mapsto \int_0^t (\mathcal{L}f)(s, X_s)\mathrm{d}s$ taking values in $\mathscr{C}$.

(ii) We recall that the function id_i solves PDE (2.3) so we have $\mathcal{L}\mathrm{id}_i = b^i$, see Sect. 2.2. Hence, taking $f = \mathrm{id}_i$ for some $i \in \{1, \ldots, d\}$ one gets $X = M^{\mathrm{id}_i} + \mathscr{A}^{\mathrm{id}_i}$, where formally

$$\mathscr{A}^{\mathrm{id}_i} = \text{``} \int_0^\cdot b^i(s, X_s)\mathrm{d}s\text{"},$$

by the first point in this Remark. Putting all components together one would get indeed

$$\mathscr{A}^{\mathrm{id}} := (\mathscr{A}^{\mathrm{id}_i})_i = \text{``} \int_0^\cdot b(s, X_s)\mathrm{d}s\text{"}.$$

Plugging this into the decomposition $\mathrm{id}(X_t) = M^{\mathrm{id}}_t + \mathscr{A}^{\mathrm{id}}_t$ and using Remark 5.10 gives the (formal) writing

$$X_t = X_0 + W^X_t + \text{``} \int_0^t b(s, X_s)\mathrm{d}s\text{"}$$

as expected. Notice, however, that in general $\mathrm{id}_i \notin \mathcal{D}_{\mathcal{L}}$ since $b \in C_T\mathcal{C}^{-\beta}$ so in general $b \notin C_T\bar{\mathcal{C}}^{0+}_c$. This is why the writing above is only formal. We will introduce an extended domain in the next section to make this argument rigorous.

6 Generalised SDEs and their Relationship with MP

In this final section we investigate the dynamics of the process X which formally solves the SDE $\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \mathrm{d}W_t$ and compare it to the solution to the martingale problem. First, we define a notion of solution for the formal SDE, a definition that amongst other things involves weak Dirichlet processes. We show that any solution to the MP is also a solution of the formal SDE and a chain rule holds (Theorem 6.5). Finally, we *close the circle* by showing that, under the stronger assumption for X to be a Dirichlet process, X being a solution to the formal SDE is equivalent to being a solution to the MP (Corollary 6.13). We recall that X is an $\mathcal{F}^X$-Dirichlet process if it is the sum of an $\mathcal{F}^X$-local martingale plus an adapted zero quadratic variation process. In this section there is always an underlying measurable space $(\Omega, \mathcal{F})$.

We make a further technical assumption on the support of the singular drift b. This assumption is a standing assumption until the end of the paper.

Assumption A2 Let $b \in C_T\bar{\mathcal{C}}_c^{(-\beta)+}$.

As mentioned above, the idea of the current section is inspired by Remark 5.12 and consists in further investigating to which extent our solution to the martingale problem is the solution of an SDE of the form

$$X_t = X_0 + W_t^X + \text{“}\int_0^t b(s, X_s)\mathrm{d}s\text{"},$$

where $X_0 \sim \mu$. We note that if $b = l$ were a function, the interpretation of “$\int_0^t l(s, X_s)\mathrm{d}s$" would indeed be the integral $\int_0^t l(s, X_s)\mathrm{d}s$. In particular, $\int_0^t l(s, X_s)\mathrm{d}s$ is well-defined for any $l \in C_T\bar{\mathcal{C}}_c^{0+}$. We will study various properties of $l \mapsto \int_0^t l(s, X_s)\mathrm{d}s$ for a reasonable class of distributions l (which includes for example $b \in C_T\bar{\mathcal{C}}_c^{(-\beta)+}$ from Assumption A2), proceeding similarly to [22].

Definition 6.1 Let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F})$. We say that a process X fulfils the *local time property* with respect to a topological vector space $B \supset C_T\bar{\mathcal{C}}_c^{0+}$ if $C_T\bar{\mathcal{C}}_c^{0+}$ is dense in B and the map from $C_T\bar{\mathcal{C}}_c^{0+}$ with values in $\mathscr{C}$ defined by

$$l \mapsto \int_0^t l(s, X_s)\mathrm{d}s$$

admits a continuous extension to B (or equivalently it is continuous with respect to the topology of B) which we denote by $A^{X,B}$.

Notice that this notion has been first defined in a different context in [22, Definition 6.1], see also [22, Remark 6.2] for the links to local time. Using the local time property we now introduce a notion of solution to SDE which is different from the martingale problem. We will then study its properties and links to the solution to the martingale problem.

Definition 6.2 Let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F})$. Given $b \in B \subset \mathcal{S}'(\mathbb{R}^d)$, we say that X is a *B-solution* to

$$X_t = X_0 + W_t + \int_0^t b(t, X_t)\mathrm{d}s,$$

if there exists a Brownian motion $W = W^X$ and

(a) X fulfils the local time property with respect to B;
(b) $b \in B$;
(c) $X_t = X_0 + W_t^X + A_t^{X,B}(b)$;
(d) X is an $\mathcal{F}^X$-weak Dirichlet process.

Remark 6.3 Some examples of B are $B = C_T\bar{\mathcal{C}}_c^{0+}$ and $B = C_T\bar{\mathcal{C}}_c^{(-\beta)+}$. Indeed, $\bar{\mathcal{C}}_c^{0+}$ is dense in $\bar{\mathcal{C}}_c^{(-\beta)+}$ since by [18, Lemma 5.4 (i)] $\mathcal{S} \subset \bar{\mathcal{C}}_c^{0+}$ and $\mathcal{S}$ is dense in $\bar{\mathcal{C}}_c^{(-\beta)+}$. Finally, by [19, Remark B.1] we conclude that $C_T\bar{\mathcal{C}}_c^{0+}$ is dense in $C_T\bar{\mathcal{C}}_c^{(-\beta)+}$.

Below we will investigate B-solutions for $B = C_T\bar{\mathcal{C}}_c^{(-\beta)+}$. We denote by

$$\mathcal{D}_{\mathcal{L}}^B := \left\{ f \in \mathcal{D}_{\mathcal{L}}^0 \text{ such that } g := \mathcal{L}f \in B \right\}. \tag{6.1}$$

Remark 6.4 Let $B = C_T\bar{\mathcal{C}}_c^{(-\beta)+}$. Notice that $f = \mathrm{id} \in \mathcal{D}_{\mathcal{L}}^B$ and $\mathcal{L}\, id = b$, in the sense that $\mathcal{L}\,\mathrm{id}_i = b^i$ for all $i = 1, \dots, d$ as recalled in Remark 5.12 item (ii).

Theorem 6.5 *Let $B = C_T\bar{\mathcal{C}}_c^{(-\beta)+}$. Let $(X, \mathbb{P})$ be the solution to the martingale problem with distributional drift b and i.c. μ. Then there exists a Brownian motion W^X with respect to $\mathbb{P}$ such that X is a B-solution of*

$$X_t = X_0 + W_t^X + \int_0^t b(s, X_s)\mathrm{d}s,$$

where $X_0 \sim \mu$. Moreover, for every $f \in \mathcal{D}_{\mathcal{L}}^B$ we have the chain rule

$$f(t, X_t) = f(0, X_0) + \int_0^t (\nabla f)(s, X_s) \cdot \mathrm{d}W_s^X + A_t^{X,B}(\mathcal{L}f), \tag{6.2}$$

and the equality

$$A_t^{X,B}(\mathcal{L}f) = \mathscr{A}_t^f. \tag{6.3}$$

Remark 6.6 Notice that point (c) in Definition 6.2 provides the standard decomposition of the weak Dirichlet process X, where the local martingale component is given by $X_0 + W^X$ and the martingale orthogonal process is given by $A_t^{X,B}(b) = \mathscr{A}_t^{\mathrm{id}}$ in view of (6.3) and Remark 6.4.

Proof of Theorem 6.5 For ease of notation we write A^X in place of $A^{X,B}$.

Let $(X, \mathbb{P})$ be the solution to the martingale problem with distributional drift b and i.c. μ. We have to show that the four conditions of Definition 6.2 are satisfied. Clearly, $b \in B$ which is point (b) of Definition 6.2. By Proposition 5.3, for every $f \in C^{0,1}$ we have that $f(t, X_t)$ is an $\mathcal{F}^X$-weak Dirichlet process; hence, X is also a weak Dirichlet process (point (d) of Definition 6.2) with decomposition

$$f(t, X_t) = M_t^f + \mathscr{A}_t^f.$$

Next we check the local time property, which is point (a) of Definition 6.2. We use that X solves the martingale problem for every $f \in \mathcal{D}_{\mathcal{L}} \subset C^{0,1}$ (thus

$f(t, X_t) - \int_0^t (\mathcal{L} f)(s, X_s) \mathrm{d}s$ is a local martingale) and uniqueness of the weak Dirichlet decomposition to get

$$\mathscr{A}^f = \int_0^{\cdot} (\mathcal{L} f)(s, X_s) \mathrm{d}s = A^X(\mathcal{L} f), \tag{6.4}$$

where the second equality holds because $\mathcal{L} f \in C_T \bar{\mathcal{C}}_c^{0+}$. We want to show that A^X extends to all $g \in B = C_T \bar{\mathcal{C}}_c^{(-\beta)+}$. Let us denote by T the map

$$\begin{aligned} T : C_T \bar{\mathcal{C}}_c^{(-\beta)+} &\to C_T D\mathcal{C}^{\beta+} \\ g &\mapsto T(g) := v, \end{aligned}$$

where v is the unique solution in $C_T \mathcal{C}^{(1+\beta)+}$ of PDE

$$\begin{cases} \mathcal{L} v = g \\ v(T) = 0, \end{cases}$$

which is PDE (2.3) with $v_T = 0$, see Sect. 2.2. It is clear that for $f \in \mathcal{D}_{\mathcal{L}}$ and $g = \mathcal{L} f \in C_T \bar{\mathcal{C}}_c^{0+}$ we have $T(g) = f$ so that (6.3) writes

$$\mathscr{A}^{T(g)} = A^X(g).$$

Now we recall that $g \mapsto T(g) \in C_T \mathcal{C}^{(1+\beta)+} \subset C^{0,1}$ is continuous, see Sect. 2.2, in particular when $g_n \to g$ in $C_T \bar{\mathcal{C}}_c^{-\beta}$ then $f_n = T(g_n) \to T(g) = f$ in $C_T \mathcal{C}^{(1+\beta)+} \subset C^{0,1}$. Moreover, by Corollary 5.9 also the map $f \mapsto \mathscr{A}^f$ is continuous from $C^{0,1}$ to $\mathscr{C}$. Now we use the density of $C_T \bar{\mathcal{C}}_c^{0+}$ in $C_T \bar{\mathcal{C}}_c^{(-\beta)+}$ to conclude that the local time property holds and also (6.3) holds. Point (c) in Definition 6.2 follows from the chain rule (6.2) (shown below) for $f = \mathrm{id}$ using Remark 6.4.

It is left to prove that the chain rule (6.2) holds. We define $W^X := M^{\mathrm{id}} - X_0$, which is a Brownian motion by Remark 5.10. First, we prove that (6.2) holds for $f \in \mathcal{D}_{\mathcal{L}}$. Indeed, by Proposition 5.11 we know that $M_t^f = f(0, X_0) + \int_0^t (\nabla f)(s, X_s) \cdot \mathrm{d}W_s^X$ so using that X is a solution to the martingale problem we easily get that (6.2) holds for $f \in \mathcal{D}_{\mathcal{L}}$. In order to extend it to $f \in \mathcal{D}_{\mathcal{L}}^B$, we use the operator T and rewrite the chain rule (6.2) as

$$(Tg)(t, X_t) - (Tg)(0, X_0) - \int_0^t \nabla(Tg)(s, X_s) \cdot \mathrm{d}W_s^X = A_t^X(g), \tag{6.5}$$

for all $g \in B = C_T \bar{\mathcal{C}}_c^{(-\beta)+}$. Notice that (6.5) holds for $g \in C_T \bar{\mathcal{C}}_c^{0+}$ since (6.2) holds for $f \in \mathcal{D}_{\mathcal{L}}$ with $\mathcal{L}\{ = g$. The left-hand side of (6.5) is continuous from B to $\mathscr{C}$ because it is the composition of continuous operators. The right-hand side of (6.5) extends from $g \in C_T \bar{\mathcal{C}}_c^{0+}$ to $g \in B$ by the local time property (a). Since $C_T \bar{\mathcal{C}}_c^{0+}$ is dense in B, then (6.5) extends to B, which is (6.2) as wanted. □

Remark 6.7 Notice that if, in the previous proof, we defined the solution operator T using a different terminal condition $v_T \in \mathcal{C}^{(1+\beta)+}$, $v_T \neq 0$, it would have led to the same operator $A^{X,B}$. This can be seen by noticing that the operator is the unique extension of the integral operator $l \mapsto \int_0^t l(s, X_s)\mathrm{d}s$.

We now introduce a refined notion of B-solution, which will be used later.

Definition 6.8 Let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F})$. We say that X is a *reinforced B-solution* of

$$X_t = X_0 + W_t^X + \int_0^t b(t, X_t)\mathrm{d}s$$

if

(i) it is a B-solution of the SDE in the sense of Definition 6.2;
(ii) for any $f \in C_b^{1,2,B}$, where

$$C_b^{1,2,B} := \{f \in C_b^{1,2} \text{ such that } \dot{f} + \frac{1}{2}\Delta f \in C_T\bar{\mathcal{C}}_c^{0+} \text{ and } \nabla f b \in B\},$$

then

$$\int_0^t (\nabla f)(s, X_s) \cdot \mathrm{d}^- A_s^{X,B}(b) = A_t^{X,B}(\nabla f\, b), \tag{6.6}$$

where the forward integral d^-A is the one given in [23] in the one-dimensional case, which can be straightforwardly extended to the vector case. In particular, for a locally bounded integrand process Y and a continuous integrator process X we denote

$$\int_0^t Y_s \cdot \mathrm{d}^- X_s = \sum_{i=1}^d \int_0^t Y_s^i \mathrm{d}^- X_s^i.$$

Remark 6.9 (i) When $b \in C_T\bar{\mathcal{C}}_c^{0+}$ and $f \in C_b^{1,2}$ then $\nabla f b \in C_T\bar{\mathcal{C}}_c^{0+}$ because we can choose the approximating sequence $b_n \to b$ with compact support to construct the approximating sequence $\nabla f b_n \to \nabla f b$. In this case equality (6.6) holds because both members are equal to $\int_0^t (\nabla f\, b)(s, X_s)\mathrm{d}s$. Thus, it is natural to require the condition (6.6).

(ii) In the case $B = C_T\bar{\mathcal{C}}_c^{(-\beta)+}$, we notice that the condition $\nabla f b \in B$ is always satisfied. Indeed, $\nabla f \in C_T\mathcal{C}^{\beta+}$ and $b \in B$; thus, by (2.2) $\nabla f b \in C_T\mathcal{C}^{(-\beta)+}$. Finally, $\nabla f b \in C_T\bar{\mathcal{C}}_c^{(-\beta)+}$ because we can construct the compactly supported sequence by considering $\nabla f b_n$, where b_n is the compactly supported sequence that converges to b in $C_T\bar{\mathcal{C}}_c^{(-\beta)+}$, using again (2.2). Thus, $C_b^{1,2,B}$ reduces to

$$\{f \in C_b^{1,2} \text{ such that } \dot{f} + \frac{1}{2}\Delta f \in C_T\bar{\mathcal{C}}_c^{0+}\}$$

and does not depend on B.

Next we want to consider the case when X is an $\mathcal{F}^X$-Dirichlet process. In this case we show that the notion of solution of the martingale problem with distributional drift is equivalent to the one of the reinforced B-solution. Let us start with a remark.

Remark 6.10 If X is a B-solution which is an $\mathcal{F}^X$-Dirichlet process, then $[X, X]_t = t\mathrm{I}_d$. Indeed, by Remark 6.6 we have that $X_t = W_t^X + A_t^{X,B}(b)$ is the standard decomposition of the weak Dirichlet process X, and by the uniqueness of the weak Dirichlet decomposition and the fact that X is an $\mathcal{F}^X$-Dirichlet process then $A_t^{X,B}(b)$ is a zero quadratic variation process and so $[X, X]_t = t\mathrm{I}_d$.

Proposition 6.11 *Let $B = C_T\bar{\mathcal{C}}_c^{(-\beta)+}$. If $(X, \mathbb{P})$ satisfies the martingale problem with distributional drift b and X is an $\mathcal{F}^X$-Dirichlet process, then X is a reinforced B-solution according to Definition 6.8.*

Proof First, we notice that point (i) of Definition 6.8 is satisfied by Theorem 6.5. Next we check point (ii) and we write A^X instead of $A^{X,B}$ for ease of notation. Let $f \in C_b^{1,2,B}$. Using the weak Dirichlet decomposition since $f \in C_b^{1,2}$, we have

$$\begin{aligned} f(t, X_t) &= M_t^f + \mathcal{A}_t^f \\ &= f(0, X_0) + \int_0^t (\nabla f)(r, X_r) \cdot \mathrm{d}W_r^X + \mathcal{A}_t^f, \end{aligned} \tag{6.7}$$

having used Proposition 5.11 to express the martingale component part.

On the other hand, it easily follows that $f \in \mathcal{D}_{\mathcal{L}}^B$ defined in (6.1), because $\mathcal{L}f = \nabla f b + g$, where $g := \dot{f} + \frac{1}{2}\Delta f \in C_T\bar{\mathcal{C}}_c^{0+} \subset B$ by assumption, and $\nabla f b \in B$ as seen in Remark 6.9, item (ii). Since X is a B-solution and an $\mathcal{F}^X$-Dirichlet process, by Remark 6.10 we have $[X, X]_t = t\mathrm{I}_d$. So, by applying a slight adaptation of Itô's formula [24, Theorem 2.2] to $f(t, X_t)$ for $f \in C_b^{1,2}$ we have

$$\begin{aligned} f(t, X_t) =& f(0, X_0) + \int_0^t (\nabla f)(r, X_r) \cdot \mathrm{d}W_r^X + \int_0^t (\nabla f)(r, X_r) \cdot \mathrm{d}^- A_r^X(b) \\ &+ \int_0^t (\partial_t f + \frac{1}{2}\Delta f)(r, X_r)\mathrm{d}r \\ =& f(0, X_0) + \int_0^t (\nabla f)(r, X_r) \cdot \mathrm{d}W_r^X + \int_0^t (\nabla f)(r, X_r) \cdot \mathrm{d}^- A_r^X(b) \\ &+ A_t^X(\partial_t f + \frac{1}{2}\Delta f). \end{aligned} \tag{6.8}$$

We recall that $\partial_t f + \frac{1}{2}\Delta f \in C_T\bar{\mathcal{C}}_c^{0+}$ since $f \in C_b^{1,2,B}$, and so $A_t^X(\partial_t f + \frac{1}{2}\Delta f)$ is trivially well-defined. On the other hand $\partial_t f + \frac{1}{2}\Delta f = \mathcal{L}f - \nabla f\, b$, where $\mathcal{L}f, \nabla f\, b \in B$ as noticed in Remark 6.9 item (ii); thus, we can write

$$A_t^X(\partial_t f + \frac{1}{2}\Delta f) = A_t^X(\mathcal{L}f - \nabla f\, b) = A_t^X(\mathcal{L}f) - A_t^X(\nabla f\, b).$$

Plugging this into (6.8) and comparing with (6.7), we get

$$\mathcal{A}_t^f = \int_0^t (\nabla f)(r, X_r) \cdot \mathrm{d}^- A_r^X \\ + A_t^X(\mathcal{L} f) - A_t^X(\nabla f\, b);$$

hence, applying (6.3) we conclude. □

The next result is the converse statement of Proposition 6.11.

Proposition 6.12 *Let $B = C_T \bar{\mathcal{C}}_c^{(-\beta)+}$ and $b \in B$. Let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F})$. Let X be a reinforced B-solution according to Definition 6.8, which is also an $\mathcal{F}^X$-Dirichlet process. Then $(X, \mathbb{P})$ solves the martingale problem with distributional drift b.*

Proof We need to show that for every $f \in \mathcal{D}_\mathcal{L}$

$$f(t, X_t) - f(0, X_0) - \int_0^t (\mathcal{L}\{)(r, X_r)\mathrm{d}r$$

is an $\mathcal{F}^X$-local martingale under $\mathbb{P}$. Since $f \in \mathcal{D}_\mathcal{L}$ we know that there exists $l \in C_T \bar{\mathcal{C}}_c^{0+}$ such that $\mathcal{L} f = l$. By the density of $\mathcal{S}$ into $\bar{\mathcal{C}}_c^{(-\beta)+}$, see [18, Lemma 5.4] and using [19, Remark B.1] we see that $C_T \mathcal{S}$ is dense in $C_T \bar{\mathcal{C}}_c^{(-\beta)+}$. Thus, we can find a sequence (b_n) such that $b_n \in C_T \bar{\mathcal{C}}_c^{0+}$ and $b_n \to b$ in $C_T \mathcal{C}^{(-\beta)+}$. Let $\mathcal{L}_n u := \partial_t u + \frac{1}{2}\Delta u + \nabla u\, b_n$ and let us consider the PDE $\mathcal{L}_n f_n = l$ and $f_n(T) = f(T)$. By [18, Remark 4.12] we know that the unique solution $f_n \in C_T \mathcal{C}^{(1+\beta)+}$ is also the classical solution as given in [21, Theorem 5.1.9]; hence, $f_n \in C_b^{1,2}$. We recall that X is a B-solution in the sense of Definition 6.2 and it is an $\mathcal{F}^X$-Dirichlet process with decomposition $X = W^X + A^{X,B}$ by Remark 6.10. By Itô's formula [25, Theorem 6.1], taking into account the linearity of $A^{X,B}$ and the fact that $b_n \in C_T \bar{\mathcal{C}}_c^{0+}$, we have

$$\begin{aligned} f_n(t, X_t) =& f_n(0, X_0) + \int_0^t (\nabla f_n)(s, X_s) \cdot \mathrm{d}W_s + \int_0^t (\nabla f_n)(s, X_s) \cdot \mathrm{d}^- A_s^{X,B}(b - b_n) \\ &+ \int_0^t (\nabla f_n)(s, X_s) b_n(s, X_s)\mathrm{d}s + \frac{1}{2}\int_0^t (\Delta f_n)(s, X_s)\mathrm{d}s + \int_0^t (\partial_s f_n)(s, X_s)\mathrm{d}s \\ =& f_n(0, X_0) + \int_0^t (\nabla f_n)(s, X_s) \cdot \mathrm{d}W_s + \int_0^t (\nabla f_n)(s, X_s) \cdot \mathrm{d}^- A_s^{X,B}(b - b_n) \\ &+ \int_0^t l(s, X_s)\mathrm{d}s, \end{aligned} \tag{6.9}$$

having used $\mathcal{L}_n f_n = l$ in the last equality. Using again the linearity of $A^{X,B}$, we have

$$\begin{aligned} &\int_0^t (\nabla f_n)(s, X_s) \cdot \mathrm{d}^- A_s^{X,B}(b - b_n) \\ &\quad = \int_0^t (\nabla f_n)(s, X_s) \cdot \mathrm{d}^- A_s^{X,B}(b) - \int_0^t (\nabla f_n)(s, X_s) \cdot \mathrm{d}^- A_s^{X,B}(b_n). \end{aligned} \tag{6.10}$$

The second integral on the RHS is equal to $A_t^{X,B}(\nabla f_n\, b_n)$ by Remark 6.9 item (i) since $f_n \in C_b^{1,2}$ and $b_n \in C_T\bar{\mathcal{C}}_c^{0+}$. Since X is a reinforced B-solution, by (6.6) the first integral on the RHS of (6.10) gives $A_t^{X,B}(\nabla f_n\, b)$ so by additivity we rewrite (6.10) as

$$\int_0^t (\nabla f_n)(s, X_s) \cdot \mathrm{d}^- A_s^{X,B}(b - b_n) = A_t^{X,B}(\nabla f_n\, b - \nabla f_n\, b_n). \tag{6.11}$$

Plugging (6.11) into (6.9), we have

$$\begin{aligned} f_n(t, X_t) - f_n(0, X_0) - A_t^{X,B}(\nabla f_n\, b - \nabla f_n\, b_n) - \int_0^t l(s, X_s)\mathrm{d}s \\ = \int_0^t (\nabla f_n)(s, X_s) \cdot \mathrm{d}W_s. \end{aligned} \tag{6.12}$$

Since $b_n \to b$ in $C_T\mathcal{C}^{-\beta}$, we then have $f_n \to f$ in $C_T\mathcal{C}^{(1+\beta)+}$ and $\nabla f_n \to \nabla f$ in $C_T\mathcal{C}^{\beta+}$ by continuity results for PDE (2.3), see Sect. 2.2. Thus, the right-hand side of (6.12) converges u.c.p. to $\int_0^t (\nabla f)(s, X_s) \cdot \mathrm{d}W_s$, which is a local martingale under $\mathbb{P}$. Moreover, the left-hand side of (6.12) converges u.c.p. to $f(t, X_t) - f(0, X_0) - \int_0^t l(s, X_s)\mathrm{d}s$ and since $l = \mathcal{L}f$ we conclude. □

As a consequence, we get a characterisation property for solutions of the SDE in terms of solutions to martingale problem.

Corollary 6.13 *Let $B = C_T\bar{\mathcal{C}}_c^{(-\beta)+}$ and $b \in B$. Let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F})$. Suppose that X is an $\mathcal{F}^X$-Dirichlet process. Then X is a reinforced B-solution of the SDE*

$$X_t = X_0 + W_t + \int_0^t b(t, X_t)\mathrm{d}t$$

if and only if $(X, \mathbb{P})$ solves the martingale problem with distributional drift b and initial condition $X_0 \sim \mu$.

Proof Combine Proposition 6.11 and Proposition 6.12. □

Acknowledgements We would like to thank the anonymous Referee for their careful reading that has led to various improvements in the final version of the paper.

Author Contributions All authors contributed equally to the manuscript.

Funding Open access funding provided by Università degli Studi di Torino within the CRUI-CARE Agreement. The research of the first named author has been partially supported by the MIUR-PRIN 2022 project "Non-Markovian dynamics and non-local equations", No. 202277N5H9. The research of the second named author has been partially supported by the ANR-22-CE40-0015-01 project (SDAIM).

Data Availability No data are associated with this article.

Conflict of interest All authors declare that they have no conflicts of interest.

Appendix A. Some Useful Results from the Literature

In this Appendix we recall a useful theorem from [21] on existence and regularity results of parabolic PDEs. Before stating the theorem, we recall the notation used in the book, see [21, Chapter 5].

The classical Hölder space $\mathcal{C}^{2+\nu}(\mathbb{R}^d)$ for $0 < \nu < 1$ was introduced in Sect. 2. For functions of two variables $(t, x) \in [0, T] \times \mathbb{R}^d$, we consider the spaces introduced in [21, Sect. 5.1]

$$\begin{aligned} C^{0,\nu}([0,T]\times\mathbb{R}^d) := \{ & f \in C([0,T]\times\mathbb{R}^d) : f(t,\cdot) \in \mathcal{C}^\nu(\mathbb{R}^d)\, \forall t \in [0,T], \\ & \sup_{t\in[0,T]} \|f(t,\cdot)\|_{\mathcal{C}^\nu} < \infty\}, \end{aligned}$$

with norm

$$\|f\|_{C^{0,\nu}([0,T]\times\mathbb{R}^d)} := \sup_{t\in[0,T]} \|f(t,\cdot)\|_{\mathcal{C}^\nu}$$

and

$$\begin{aligned} & C^{1,2+\nu}([0,T]\times\mathbb{R}^d) \\ & := \{f \in C^{1,2}([0,T]\times\mathbb{R}^d) : \partial_t f, \partial_{x_i,x_j} f \in C^{0,\nu}([0,T]\times\mathbb{R}^d),\ \forall i,j = 1,\dots,d\} \end{aligned}$$

with norm

$$\|f\|_{C^{1,2+\nu}([0,T]\times\mathbb{R}^d)} := \|f\|_\infty + \sum_{i=1}^d \|\partial_i f\|_\infty + \|\partial_t f\|_\infty + \sum_{i,j=1}^d \|\partial_{ij} f\|_{C^{0,\nu}([0,T]\times\mathbb{R}^d)}.$$

Remark A.1 Note that the following is an equivalent norm in $\mathcal{C}^\nu$

$$\sup_x |f(x)| + \sup_{x_1,x_2\, x_1\neq x_2} \frac{|f(x_1) - f(x_2)|}{|x_1 - x_2|^\nu},$$

namely we can freely choose not to restrict x_1, x_2 to a bounded interval.

Remark A.2 Note that if $f \in C_T\mathcal{C}^\nu$ then trivially we have $f \in C^{0,\nu}([0,T]\times\mathbb{R}^d)$ and if $f \in C^{1,2+\nu}([0,T]\times\mathbb{R}^d)$ then trivially $f \in C^{1,2}_{buc}$.

Let $a_{i,j}, b_i, c, f : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ be uniformly continuous and belonging to $C^{0,\nu}([0,T] \times \mathbb{R}^d)$ with $0 < \nu < 1$. Let a satisfy the uniform ellipticity condition $\sum_{i,j=1}^d a_{i,j}(t,x)\xi_i\xi_j \geq \lambda|\xi|^2$ for $t \in [0,T],\ x, \xi \in \mathbb{R}^d$, for some $\lambda > 0$. Let $u_0 \in \mathcal{C}^{2+\nu}(\mathbb{R}^d)$.

Let us consider the second-order operator

$$\mathcal{A}(t,x) = \sum_{i,j=1}^d a_{i,j}(t,x)\partial_{x_i x_j} + \sum_{i=1}^d b_i(t,x)\partial_{x_i} + c(t,x)$$

and the PDE

$$\begin{cases} \partial_t u(t,x) = \mathcal{A}(t,x)u(t,x) + f(t,x),\ (t,x) \in [0,T] \times \mathbb{R}^d \\ u(0,x) = u_0(x),\ x \in \mathbb{R}^d. \end{cases} \tag{A.1}$$

Theorem A.3 *(Theorem 5.1.9 in [21]) Let $a_{i,j}, b_i, c, f, u_0$ be as above. Then PDE* (A.1) *has a unique solution $u \in C^{1,2+\nu}([0,T] \times \mathbb{R}^d)$ and*

$$\|u\|_{C^{1,2+\nu}([0,T]\times\mathbb{R}^d)} \leq C(\|u_0\|_{\mathcal{C}^{2+\nu}(\mathbb{R}^d)} + \|f\|_{C^{0,\nu}([0,T]\times\mathbb{R}^d)}),$$

for some $C > 0$.

References

1. Athreya, S., Butkovsky, O., Mytnik, L.: Strong existence and uniqueness for stable stochastic differential equations with distributional drift. Ann. Probab. **48**(1), 178–210 (2020)
2. Bahouri, H., Chemin, J.-Y., Danchin, R.: Fourier Analysis and Nonlinear Partial Differential Equations. Springer, Berlin (2011)
3. Barlow, M.T., Yor, M.: Semi-martingale inequalities via the Garsia-Rodemich-Rumsey lemma, and applications to local times. J. Funct. Anal. **49**(2), 198–229 (1982)
4. Bony, J.-M.: Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires. Ann. Sci. Ec. Norm. Super. **14**, 209–246 (1981)
5. Cannizzaro, G., Chouk, K.: Multidimensional SDEs with singular drift and universal construction of the polymer measure with white noise potential. Ann. Probab. **46**(3), 1710–1763 (2018)
6. Chaudru de Raynal, P.-E., Menozzi, S.: On multidimensional stable-driven stochastic differential equations with besov drift. Electron. J. Probab. **27**, 1–52 (2022)
7. Delarue, F., Diel, R.: Rough paths and 1d SDE with a time dependent distributional drift: application to polymers. Probab. Theory Related Fields **165**(1–2), 1–63 (2016)
8. Dellacherie, C., Meyer, P.-A.: Probabilities and potential. Transl. from the French, volume 29 of North-Holland Math. Stud. Elsevier, Amsterdam, (1978)
9. Dunford, N., Schwartz, J.T.: Linear operators. Part I. Wiley Classics Library. John Wiley & Sons Inc., New York, General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication (1988)
10. Errami, M., Russo, F.: Covariation de convolutions de martingales. C. R. Acad. Sci. Paris Sér. I Math. **326**(5), 601–606 (1998)
11. Errami, M., Russo, F.: n-covariation, generalized Dirichlet processes and calculus with respect to finite cubic variation processes. Stoch. Process. Appl. **104**(2), 259–299 (2003)
12. Flandoli, F., Issoglio, E., Russo, F.: Multidimensional SDEs with distributional coefficients. T. Am. Math. Soc. **369**, 1665–1688 (2017)
13. Flandoli, F., Russo, F., Wolf, J.: Some SDEs with distributional drift. I. General calculus. Osaka J. Math. **40**(2), 493–542 (2003)

14. Flandoli, F., Russo, F., Wolf, J.: Some SDEs with distributional drift. II. Lyons-Zheng structure, Itô's formula and semimartingale characterization. Random Oper. Stoch. Equ. **12**(2), 145–184 (2004)
15. Gozzi, F., Russo, F.: Weak Dirichlet processes with a stochastic control perspective. Stoch. Process. Appl. **116**(11), 1563–1583 (2006)
16. Gubinelli, M., Imkeller, P., Perkowski, N.: Paracontrolled distributions and singular PDEs. Forum Math. Pi **3**, 75 (2015)
17. Issoglio, E., Pagliarani, S., Russo, F., Trevisani, D.: Degenerate McKean-Vlasov equations with drift in anisotropic negative Besov spaces. Preprint arxiv:2401.09165, (2014)
18. Issoglio, E., Russo, F.: A PDE with drift of negative Besov index and linear growth solutions. Differ. Integral Equ. **37**(9–10), 585–622 (2024)
19. Issoglio, E., Russo, F.: McKean SDEs with singular coefficients. Ann. l'Institut Henri Poincaré Probabilités et Statistiques. **59**(3), 1530–1548 (2023)
20. Karatzas, I., Shreve, S.E.: Brownian Motion and Stochastic Calculus. Graduate Texts in Mathematics, Springer, New York (1991)
21. Lunardi, A.: Analytic semigroups and optimal regularity in parabolic problems. Progress in Nonlinear Differential Equations and their Applications, 16. Birkhäuser Verlag, Basel, 1995
22. Russo, F., Trutnau, G.: Some parabolic PDEs whose drift is an irregular random noise in space. Ann. Probab. **35**(6), 2213–2262 (2007)
23. Russo, F., Vallois, P.: Forward, backward and symmetric stochastic integration. Probab. Theory Related Fields **97**(3), 403–421 (1993)
24. Russo, F., Vallois, P.: The generalized covariation process and Itô formula. Stoch. Process. Appl. **59**(1), 81–104 (1995)
25. Russo, F., Vallois, P.: Stochastic calculus with respect to continuous finite quadratic variation processes. Stoch. Stoch. Rep. **70**(1–2), 1–40 (2000)
26. Russo, F., Vallois, P.: Stochastic Calculus via Regularizations, vol. 11. Springer International Publishing, Springer-Bocconi (2022)
27. Seignourel, P.: Processus dans un milieu irrégulier: une approche par les formes de Dirichlet. École polytechnique, 1998
28. Stroock, D.W., Varadhan, S.R.S.: Multidimensional diffusion processes. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 233. Springer-Verlag, Berlin (1979)
29. Zhang, X., Zhao, G.: Heat kernel and ergodicity of SDEs with distributional drifts. Preprint arxiv:1710.10537, (2017)
30. Zvonkin, A.K.: A transformation of the phase space of a diffusion process that removes the drift. Math. USSR Sb **22**, 129–149 (1975)