The Annals of Probability
2020, Vol. 48, No. 1, 178–210
https://doi.org/10.1214/19-AOP1358

STRONG EXISTENCE AND UNIQUENESS FOR STABLE STOCHASTIC DIFFERENTIAL EQUATIONS WITH DISTRIBUTIONAL DRIFT

BY SIVA ATHREYA[1], OLEG BUTKOVSKY[2] AND LEONID MYTNIK[3]

[1] *Indian Statistical Institute, athreya@isibang.ac.in*
[2] *Weierstrass Institute, oleg.butkovskiy@gmail.com*
[3] *Faculty of Industrial Engineering and Management, Technion—Israel Institute of Technology, leonid@ie.technion.ac.il*

We consider the stochastic differential equation

$$dX_t = b(X_t)\,dt + dL_t,$$

where the drift b is a generalized function and L is a symmetric one dimensional α-stable Lévy processes, $\alpha \in (1,2)$. We define the notion of solution to this equation and establish strong existence and uniqueness whenever b belongs to the Besov–Hölder space $\mathcal{C}^\beta$ for $\beta > 1/2 - \alpha/2$.

1. Introduction. In this article, we consider the stochastic differential equation (SDE)

$$X_t = x + \int_0^t b(X_s)\,ds + L_t, \quad t \ge 0, \tag{1.1}$$

where the initial condition $x \in \mathbb{R}$, L is a symmetric 1-dimensional α-stable process with $\alpha \in (1,2)$ (i.e., a Markov process whose generator is given by (2.3)), and the drift b is in the Hölder–Besov space $\mathcal{C}^\beta = \mathcal{C}^\beta(\mathbb{R},\mathbb{R})$ for $\beta \in \mathbb{R}$ (see [37], Definition 7). When $\beta \le 0$, this equation is not well posed in the classical sense. Indeed, in this case b is not a function but just a distribution and the expression $b(X_s)$ is not well defined. Thus it is not clear a priori what should be called a *solution* to the SDE. Inspired by the Bass–Chen approach [6], we formulate a natural notion of a solution to (1.1) (see Definition 2.1) and establish strong existence and pathwise uniqueness of a solution when $\beta > \frac{1-\alpha}{2}$; see Theorem 2.3.

It has been well known for quite a long time that ordinary differential equations (ODEs) regularize when an additional forcing by Brownian motion is added. Indeed, if $b\colon \mathbb{R}^d \to \mathbb{R}^d$, $d \ge 1$, is a β-Hölder function, $0 < \beta < 1$, then an ODE

$$dX_t = b(X_t)\,dt, \quad t \ge 0$$

might have multiple solutions or no solutions when b is a bounded measurable function. However, once the random forcing by Brownian motion $(B_t)_{t\ge 0}$ is added, the corresponding SDE

$$dX_t = b(X_t)\,dt + dB_t, \quad t \ge 0 \tag{1.2}$$

has a unique strong solution even for bounded measurable b without any additional assumptions on continuity. This phenomenon is called "regularization by noise" in the literature. For SDE (1.2), the strong existence and uniqueness of solutions was established by Zvonkin in [46] in the case $d = 1$ and extended by Veretennikov [44] to the multidimensional setting. Later Krylov and Röckner [29] generalized this result for the case of a locally unbounded b under a suitable integrability condition. In all the above cases, the proofs use a Zvonkin-type transformation [46] that allows to make the "nonregular" drift much more regular.

Received January 2018; revised November 2018.
MSC2010 subject classifications. 60H10, 60G52.
Key words and phrases. Stochastic differential equations, strong solution, regularization by noise, stable processes, Zvonkin transformation.

It turns out that it is possible to consider drifts that are not functions but generalized functions. In this case, one needs to specify what exactly is meant by a solution to (1.2) since it is not straightforward to define the term $\int_0^t b(X_s)\,ds$. Among the first works studying the question of weak existence and uniqueness for SDEs with generalized drift, we would like to mention [38, 39] by Portenko. Further results were obtained by Harisson and Shepp [26] who showed strong existence and uniqueness of solutions to (1.2) in $d = 1$ if $b(\cdot)$ equals a constant c times the delta function and $|c| < 1$. Le Gall [34] generalized this result for the case where b is a finite signed measure.

A general approach for studying SDEs with distributional drift was developed by Bass and Chen in [6]. They suggested a natural definition of a solution to SDE (1.2) via an approximating scheme and established strong existence and uniqueness for (1.2) whenever b is the distributional derivative of $\mathcal{C}^\gamma$ functions with $\gamma > 1/2$. Their main tool was again the Zvonkin method; they used the fact that for $d = 1$ the Zvonkin transformation can completely eliminate the drift.

The above question has also been studied for other types of forcing instead of Brownian motion. For the case where SDE is driven by a fractional Brownian motion, we refer the reader to [5]; the results about general continuous forcings can be found in [11]. In the case of a forcing by a pure jump process, it is clear that this process should have "sufficiently many" small jumps. That is, if Brownian motion is replaced in (1.2) just by the standard Poisson process, then this does not give any improvement in the regularity properties of the equation. Indeed, the equation will already have multiple solutions while still "waiting" for the first jump of the Poisson process. Thus it is natural to expect that the bigger the intensity of small jumps the rougher drift b can be.

Indeed, Tanaka, Tsuchiya, Watanabe in [42] proved that in the case $d = 1$ equation (1.1) has a pathwise unique solution if L is a symmetric α-stable process, b is a bounded Borel measurable function and $\alpha > 1$ (recall that the bigger the parameter $\alpha \in (0, 2)$, the higher is the intensity of small jumps). On the other hand, it was also shown in [42] that if $\alpha \in (0, 1)$ and b is bounded Hölder continuous with exponent β, where $0 < \beta < 1 - \alpha$, then equation (1.1) might have multiple solutions. The case of higher dimensions was resolved by Priola in [40] who showed that in the case of dimension $d \geq 2$ and $\alpha \in [1, 2)$, the pathwise uniqueness holds for this equation if the drift b is in $\mathcal{C}^\beta$ and $\beta > 1 - \alpha/2$. This result was extended by Chen, Song and Zhang in [12] to the case $\alpha \in (0, 1)$.

Let us also mention other related works. Bogachev and Pilipenko in [10] showed strong existence and uniqueness for (1.1) for b being a function of bounded variation and belonging to a certain Kato class (see [10], Definition 1, (9) and (10)). Malliavin differentiability of strong solutions to SDEs driven by truncated α-stable process and with drift $b \in \mathcal{C}^\beta$, $\beta > 2 - \alpha$, $\alpha \in (1, 2)$ was shown in [24].

From the discussion above, the reader may notice the following gap between the cases of $\alpha < 2$ and $\alpha = 2$. For $\alpha \in (1, 2)$, the pathwise uniqueness for (1.1) if $d = 1$ is shown by Tanaka, Tsuchiya and Watanabe in [42] for b being a bounded Borel measurable function. However, in the case of $\alpha = 2$ and $d = 1$ Bass and Chen [6] have shown that the strong existence and uniqueness hold under much milder assumptions on b; namely b can be the distributional derivative of $\mathcal{C}^\gamma$ functions with $\gamma > 1/2$. This paper closes this gap, by showing that for $\alpha \in (1, 2)$ in dimension $d = 1$ the strong existence and uniqueness hold for (1.1) under much more relaxed conditions on the drift b than in [42]. Our main result in Theorem 2.3 states that for $\alpha \in (1, 2)$ there is a unique strong solution to (1.1) if b is in the Hölder–Besov space $\mathcal{C}^\beta$ for $\beta > \frac{1-\alpha}{2}$. That is, loosely speaking, b is allowed to be a distributional derivative of a Hölder continuous function with the Hölder exponent greater than $\frac{3-\alpha}{2}$. We note that this bound on the regularity of b exactly matches the result of Bass and Chen [6] for the case

$\alpha = 2$. To the best of our knowledge, our result is the first strong existence and uniqueness result for stable SDEs with a general distributional drift.

In this article, we mainly consider *strong* solutions to (1.1). Let us briefly mention that other notions of existence and uniqueness have also been studied for (1.2) and (1.1). Weak existence and uniqueness results for (1.2) with generalized drifts have been obtained in [7], [19], [20], [18] and [45]. The question of weak uniqueness and existence for (1.1) was studied in Kulik [30] for b measurable and locally bounded and in Kim and Song [28], Chen and Wang [13] for b from a certain Kato class. Some of these results are also valid for the case when the SDEs have a nontrivial diffusion coefficient. A stronger notion of path-by-path uniqueness has been established for (1.2) in a seminal work by Davie [15] when b is a bounded measurable function and it has been generalized by Priola [41] for (1.1) with $\alpha \in (0, 2)$ and b is a bounded continuous function in $\mathcal{C}^\beta$ with $\beta > 1 - \frac{\alpha}{2}$.

In the next section, we present the main result of the paper.

2. Main result and overview of proof. We begin with introducing the basic notation and definitions. For $k \in \mathbb{Z}_+$, $D \subset \mathbb{R}^k$ and function $f : D \to \mathbb{R}$, we denote its supremum norm by $\|f\| := \sup_{x\in D} |f(x)|$. If the function f is random, then supremum in the definition of $\|f\|$ will be taken only over *nonrandom* variables. For $f, g : \mathbb{R} \to \mathbb{R}$, we define $\langle f, g\rangle := \int_{\mathbb{R}} f(x)g(x)\,dx$.

We denote by $\mathcal{C}_b^\infty$ the space of all bounded and infinitely differentiable functions $\mathbb{R} \to \mathbb{R}$. Let $\mathcal{C}_c^\infty$ be the space of all functions from $\mathcal{C}_b^\infty$ with compact support. Let $\mathcal{S}$ be the space of Schwartz functions $\mathbb{R} \to \mathbb{R}$ and let $\mathcal{S}'$ be its dual space, that is, the space of Schwartz distributions. We will work with the Besov–Hölder spaces $\mathcal{C}^\gamma := \mathcal{B}^\gamma_{\infty,\infty}$, where $\gamma \in \mathbb{R}$, which are defined using the Littlewood–Paley blocks (see the Supplementary Material [3], Section 1, for a precise definition). Let $\|\cdot\|_\gamma$ be the norm associated with the space $\mathcal{C}^\gamma$, $\gamma \in \mathbb{R}$.

Recall that for $\gamma \in (0, \infty) \setminus \mathbb{N}$ the space $\mathcal{C}^\gamma$ is the usual Hölder space of functions that are $\lfloor\gamma\rfloor$ times continuously differentiable and whose $\lfloor\gamma\rfloor$-th derivative is Hölder continuous with exponent $\gamma - \lfloor\gamma\rfloor$. For $\gamma \in (-1, 0)$ the space $\mathcal{C}^\gamma$ includes all derivatives (in the distributional sense) of Hölder functions with exponent $\gamma + 1$.

Let $\gamma \in \mathbb{R}$. In what follows, we say that a sequence of functions $(f_n)_{n\in\mathbb{Z}_+}$ converges to a function f in $\mathcal{C}^{\gamma-}$ as $n \to \infty$ if there exists $N \in \mathbb{Z}_+$ such that $\sup_{n\ge N} \|f_n\|_\gamma \le 2\|f\|_\gamma$ and

$$\lim_{n\to\infty} \|f_n - f\|_\eta = 0 \quad \text{for any } \eta < \gamma.$$

It is well known (see [43] for a detailed introduction of Besov spaces or [25], Theorem 3.23) that for any $f \in \mathcal{C}^\gamma$, there is a sequence of functions $(f_n)_{n\in\mathbb{Z}_+} \subset \mathcal{C}_b^\infty$ such that $f_n \to f$ in $\mathcal{C}^{\gamma-}$ as $n \to \infty$.

2.1. *Main result and discussion.* In this article, we study stochastic differential equation (1.1). First, let us fix a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}, \mathsf{P})$ on which $L = (L_t)_{t\ge0}$ is a symmetric 1-dimensional α-stable process, for $\alpha \in (1, 2)$. Furthermore, $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}, \mathsf{P})$ is sufficiently rich to contain other processes considered in this paper. Recall that since the drift b is not a function but just a distribution, is not straightforward to define the notion of the solution to this equation. Inspired by [6], Definition 2.1, we give the following definition.

DEFINITION 2.1. Let $\beta \in \mathbb{R}$, $b \in \mathcal{C}^\beta$, $\alpha \in (1, 2)$ and $L = (L_t)_{t\ge0}$ be a symmetric 1-dimensional α-stable process. We say that a càdlàg process $X = (X_t)_{t\ge0}$ is a solution to (1.1) with the initial condition $x \in \mathbb{R}$ if there exists a continuous process $A = (A_t)_{t\ge0}$ such that:

1. $X_t = x + A_t + L_t$, $t \ge 0$;

2. for any sequence of functions $(b_n)_{n\in\mathbb{Z}_+}$ such that $b_n \in \mathcal{C}_b^\infty$, $n \in \mathbb{Z}_+$ and $b_n \to b$ in $\mathcal{C}^{\beta-}$, as $n \to \infty$ we have

$$\mathcal{A}_t^n := \int_0^t b_n(X_s)\,ds \to A_t \quad \text{as } n \to \infty \tag{2.1}$$

in probability uniformly over bounded time intervals.

A similar definition for the Brownian case is also stated in [45], Definition 3.9. Note that for $\beta > 0$, Definition 2.1 coincides with the standard definition of a solution.

DEFINITION 2.2. Let $\rho \in (0, 1]$. We say that a solution X to (1.1) belongs to class $\mathcal{V}(\rho)$ if for any $T > 0$ and any $\kappa < \rho$ there exists $C = C(T, \kappa) > 0$ such that

$$\mathsf{E}|A_t - A_s|^2 \le C|t-s|^{2\kappa}, \quad s, t \in [0, T]. \tag{2.2}$$

Given a symmetric α-stable process L on a probability space, a strong solution to (1.1) is a càdlàg process X that is adapted to the complete filtration generated by L and which is a solution to (1.1). A weak solution of (1.1) is a couple (X, L) on the complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}, \mathsf{P})$ such that X_t is adapted to $\mathcal{F}_t$, L_t is an $(\mathcal{F}_t)_{t\ge0}$ adapted symmetric α-stable process and X is a solution to (1.1). We say weak uniqueness holds for (1.1) if whenever (X, L) and $(\widetilde{X}, \widetilde{L})$ are two weak solutions of (1.1) and X_0 has the same distribution as $\widetilde{X}_0$, then the process $(X_t)_{t\ge0}$ has the same law as the process $(\widetilde{X}_t)_{t\ge0}$. We say pathwise uniqueness holds for (1.1) if whenever (X, L) and $(\widetilde{X}, L)$ are two weak solutions of (1.1) with common L on a common probability space (w.r.t. possibly different filtrations) and with the same initial condition, then $\mathsf{P}(X_t = \widetilde{X}_t \text{ for all } t \ge 0) = 1$. We say that strong uniqueness holds for (1.1) if whenever X and $\widetilde{X}$ are two strong solutions of (1.1) relative to L with the common initial condition X_0, then $\mathsf{P}(X_t = \widetilde{X}_t \text{ for all } t \ge 0) = 1$. Clearly, pathwise uniqueness implies strong uniqueness.

As mentioned earlier, pathwise uniqueness for SDE (1.1) when $\beta > 0$ was established by Tanaka, Tsuchiya, Watanabe [42]. We present now our main theorem, which extends the result of [42] for the case $\beta \in (\frac{1-\alpha}{2}, 0]$.

THEOREM 2.3. *For any $x \in \mathbb{R}$, $\alpha \in (1, 2)$, $\beta > \frac{1-\alpha}{2}$ and $b \in \mathcal{C}^\beta$, stochastic differential equation* (1.1) *has a unique strong solution in the class* $\mathcal{V}((1 + \frac{\beta}{\alpha}) \wedge 1)$.

REMARK 2.4. We note that in our setting requirement 2 of Definition 2.1 follows from a weaker condition. Let $X = x + A + L$ be a càdlàg adapted process in class $\mathcal{V}((1 + \frac{\beta}{\alpha}) \wedge 1)$ and assume that (2.1) holds only for *a particular* sequence of smooth functions $(\widetilde{b}_n)_{n\in\mathbb{Z}_+}$ converging to b in $\mathcal{C}^{\beta-}$. Then, under the hypothesis of Theorem 2.3, it follows from our proofs (see the proof of Proposition 2.8) that (2.1) holds also for *any other* smooth sequence $(b_n)_{n\in\mathbb{Z}_+}$ converging to b in $\mathcal{C}^{\beta-}$.

Though we do not consider the case $\alpha = 2$ (i.e., when L is replaced by the standard Brownian motion B), our proof can be suitably modified to show that Theorem 2.3 holds for case $\alpha = 2$ (with an appropriate replacement of L by B in Definition 2.1). We note that in this case the result would be less restrictive than the corresponding result in [6]. Indeed, we allow $(b_n)_{n\in\mathbb{Z}_+}$ to be an arbitrary sequence of smooth functions approximating b, whereas [6] imposes extra conditions on regularity of $\mathcal{A}^n$ (cf. Definition 2.1 and [6], condition (iii) of Definition 2.1). In a recent preprint [45], weak solutions for equation (1.2) in the multidimensional case are considered. In Definition 3.9 of [45], the sequence $(b_n)_{n\in\mathbb{Z}_+}$ is allowed to be an arbitrary sequence of twice continuously differentiable functions approximating b in a suitable normed space.

To prove our result, we further develop the Zvonkin drift transformation method discussed in the Introduction. There are a number of key differences in comparison to the Brownian case considered in [6]. First, the Zvonkin transformation cannot be written down explicitly. Second, this transformation does not eliminate the drift entirely. Third, the proof of the necessary Krylov type estimate is not straightforward, since X does not have the required moments of high order. Therefore, one cannot imitate the techniques from [6] directly. An additional challenge for proving the main result comes from the following observation. Under assumptions of Theorem 2.3, if $\beta < 0$ and X is in class $\mathcal{V}((1+\frac{\beta}{\alpha}) \wedge 1)$, then the upper bound on κ from Definition 2.2 is less than 1, and thus the process A_t could be of infinite variation. However, since (2.2) holds with $\kappa > 1/2$, A has zero quadratic variation. This makes X a Dirichlet process, but not necessarily a semimartingale.

We conclude this subsection with some remarks on possible research directions. One could suitably reformulate (1.1) in higher dimensions and try to see whether existence and uniqueness (weak or strong) hold. As we will see later in Section 5, where we discuss existence of weak solutions in $d = 1$, it is easy to adapt our arguments to show weak existence in any dimension. Further, by a suitable adaptation of the arguments presented in [45] it may be possible to derive the weak uniqueness as well. As for the pathwise uniqueness in dimensions $d \geq 2$, we would like to note that, under the current assumptions on β and α, our argument does not work (see Remark 2.11). It should be further noted that even deriving weak existence in $d \geq 2$ would require a much deeper understanding; in particular, generalizations of many regularity estimates from Section 4 will be needed. Since our focus is on showing the strong existence/uniqueness, we do not discuss higher dimensions in this paper in order not to deviate attention of the reader from the key ideas.

Another interesting direction is to try to give an alternative definition of the solution via local time, that is, formally speaking, representing the solution of (1.1) as

$$X_t = x_0 + \int_{\mathbb{R}} b(x) l_t^x \, dx + L_t, \quad t \geq 0,$$

where l_t^x is the local time of X at x up to time t. It is not difficult to show that the local time l_t^x exists (see Remark 5.4). However, in order to define the integral $\int_{\mathbb{R}} b(x) l_t^x \, dx$ rigorously, one has to show that $x \mapsto l_t^x$ falls in the right regularity class. This is very challenging and will be subject of our future work.

The next subsection is devoted to the overview of the proof of the main result.

2.2. *Overview of the proof of Theorem* 2.3.

The proof of Theorem 2.3 consists of two parts: namely, existence and uniqueness. Usually proving existence is an "easy" part of these types of theorems. Indeed, in the case when the coefficients in the stochastic differential equation are sufficiently regular, it is possible to directly show strong existence. For example, for equation (1.1) when b is a bounded continuous function strong existence follows via a simple compactness argument (see the comment before [40], Lemma 4.1). However, for the equations with a generalized drift the situation is much more complicated, since even the notion of a solution should be defined very carefully.

In the intermediate steps of our proof, we will use additionally the notion of a *virtual solution*, which has been introduced recently for related equations with distributional drift (see [18], Definition 25). The notion is based on applying a Zvonkin-type transformation to the equation of interest and obtaining a "transformed" equation where the drift is more regular (see [46] and [44]). The broad strategy of the proof then involves showing existence and uniqueness for the "transformed" equation; its solution is called a "virtual solution."

However, it is not obvious at all how to identify the concept of a solution to (1.2) with the virtual solution. Recently, it was shown in [45] for some multidimensional equations driven by the Brownian motions that the virtual solutions are solutions to the martingale problem associated with the original equation. It is technically challenging to carry out the program for proving Theorem 2.3; in particular, as mentioned before, the solution $X = (X_t)_{t\geq 0}$ will not be a semimartingale. So the classical tools and methods will not be applicable. The novelty of our approach is in working with the correct notion of a (natural) solution, a suitable adaptation of the transformation, along with a specific technique for the identification of solutions and virtual solutions.

ASSUMPTION 2.5. For the rest of the paper, we fix $\alpha \in (1, 2)$, $\beta \in (\frac{1-\alpha}{2}, 0)$, $b \in \mathcal{C}^\beta$, the initial condition $x \in \mathbb{R}$ and the length of the time interval $T > 0$.

Our goal is to show that for the parameters chosen above, the stochastic differential equation (1.1) has a unique strong solution in the time interval $[0, T]$. Note also that we do not lose generality by choosing $\beta < 0$, since if we prove existence and uniqueness for any $b \in \mathcal{C}^{\beta_0}$, then it also holds for any $b \in \mathcal{C}^\beta$ with $\beta \geq \beta_0$.

As mentioned earlier, to study (1.1) we use a new version of the Zvonkin method. With the help of a certain auxiliary function $u\colon \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$, one can consider the process $Y_t := u(t, X_t)$, $t \geq 0$. This process satisfies a new SDE (which we will call the Zvonkin equation) with better drift and worse (though still not "too bad") noise. If one can prove that this Zvonkin SDE has a unique strong solution and the function u is "nice," then this would imply that the original SDE (1.1) also has a unique strong solution.

Thus everything depends on the choice of the transformation function u. In the original papers [46] and [44], the function u was a solution of a certain parabolic partial differential equation. Motivated by [17], Priola [40] suggested to take a different u, which arises from a family of resolvent equations. We further develop Priola's approach to accommodate distributional drift.

To present the equation on u, we need to recall a couple of notions. In what follows, $\mathcal{L}_\alpha$ will denote the fractional Laplace operator $-(-\Delta)^{\alpha/2}$, and $\mathcal{D}(\mathcal{L}_\alpha)$ its domain. Recall that $\mathcal{S} \subset \mathcal{D}(\mathcal{L}_\alpha)$ and if $f \in \mathcal{S}$, then

$$\mathcal{L}_\alpha f(x) = \int_{\mathbb{R}} \big(f(x+y) - f(x) - yf'(x)\mathbb{1}_{|y|\leq 1}\big)c_\alpha |y|^{-1-\alpha}\, dy, \tag{2.3}$$

with $c_\alpha > 0$ (see [33], (2.4) and Definition 2.6).

DEFINITION 2.6. Let $\{P_t\}_{t\geq 0}$ be the Markov semigroup with the infinitesimal generator $\mathcal{L}_\alpha$.

Note that $\mathcal{L}_\alpha$ is the generator of the symmetric one-dimensional α-stable process L. We extend the definition of $\mathcal{L}_\alpha$ to the space of all Schwarz distributions in the standard way. Namely, for $f \in \mathcal{S}'$ we set

$$\langle \mathcal{L}_\alpha f, \phi\rangle := \langle f, \mathcal{L}_\alpha \phi\rangle, \quad \phi \in \mathcal{S}.$$

We will be also dealing with products of a function and a distribution. In this regard, let us recall that if $f \in \mathcal{C}^{\gamma_1}$ and $g \in \mathcal{C}^{\gamma_2}$, where $\gamma_1, \gamma_2 \in \mathbb{R}$ and $\gamma_1 + \gamma_2 > 0$, then the product fg is well defined as a distribution. More precisely, the map $(f, g) \to fg$ extends to a continuous bilinear map from $\mathcal{C}^{\gamma_1} \times \mathcal{C}^{\gamma_2} \to \mathcal{C}^{\gamma_1 \wedge \gamma_2}$; see, for example, [23], Corollary 1.

Now we can present the equation on the transformation function u. We consider the following equation:

$$\lambda u - \mathcal{L}_\alpha u - f u' = g, \tag{2.4}$$

where $\lambda > 0$, $f, g \in \mathcal{C}^\eta$, $\eta \in \mathbb{R}$. We understand this equation in the distributional sense: we say that $u \in \mathcal{C}^\gamma$ is a solution to (2.4) if $\gamma > 1 - \eta$ and for any $\phi \in \mathcal{S}$,

$$\langle \lambda u - \mathcal{L}_\alpha u - f u', \phi \rangle = \langle g, \phi \rangle.$$

We note that the term fu' above involves the product of a function and a distribution. However, thanks to our additional assumption $\gamma > 1 - \eta$ and the explanations above, this product is well defined.

Clearly, for $\eta > 0$ and $\gamma > \alpha$, equation (2.4) can be interpreted pointwise.

We will call (2.4) *the resolvent equation* and we are going to use it extensively throughout the proof. In different parts of the proof, we will be substituting f and g by the drift b, its smooth approximations b_n or sometimes just by 0. For brevity, we will say $u^\lambda_{f,g}$ solves (2.4) to imply that $u^\lambda_{f,g}$ is a solution to (2.4) with the parameters λ, f and g. However, if it is clear from the context, we may drop the additional indices.

Note that the difference of our approach and [40] is that we allow f and g in (2.4) to be distributions (and not just regular functions). It will make establishing corresponding estimates much more trickier; on the other hand, it will allow us to deal with the distributional drift in our main SDE (1.1).

Our first step is to show that the resolvent equation (2.4) is actually well defined. That is, it has a unique solution with prescribed regularity and possesses a continuity property.

PROPOSITION 2.7. *For any $\eta > \frac{1}{2} - \frac{\alpha}{2}$ and $M > 0$, there exists $\lambda_0 = \lambda_0(\eta, M)$ such that for any $\lambda \geq \lambda_0$ and any $f, g \in \mathcal{C}^\eta$ with $\|f\|_\eta \leq M$ the following hold*:

(i) *there exists a unique solution $u^\lambda_{f,g}$ to (2.4) in $\mathcal{C}^{\frac{1+\alpha}{2}}$. Furthermore, for each $\gamma \in [0 \vee \eta, \alpha + \eta)$ we have $u^\lambda_{f,g} \in \mathcal{C}^\gamma$ and there exists a constant $C = C(\eta, \gamma) > 0$ such that*

$$\|u^\lambda_{f,g}\|_\gamma \leq C\lambda^{-1-\frac{\eta}{\alpha}+\frac{\gamma}{\alpha}} \|g\|_\eta (1 + \|f\|_\eta); \tag{2.5}$$

(ii) *for any sequences of functions $(f_n)_{n\in\mathbb{Z}_+}$, $(g_n)_{n\in\mathbb{Z}_+}$ such that $f_n \to f$ and $g_n \to g$ in $\mathcal{C}^{\eta-}$ as $n \to \infty$, we have*

$$\|u^\lambda_{f_n,g_n} - u^\lambda_{f,g}\|_{(1+\alpha)/2} \to 0 \quad \text{as } n \to \infty.$$

Our next step is to derive the Zvonkin equation, which is more challenging in our case due to the fact that b is a distribution. Let $u^\lambda_b = u^\lambda_{b,b}$ be the unique solution to (2.4), which exists by Proposition 2.7. We would like to apply the Zvonkin-type transform $\phi(x) = x + u^\lambda_b(x)$ to X solving (1.1). As mentioned earlier, a solution to SDE (1.1) is not a semimartingale. Thus we cannot use the standard Itô formula and have to employ and develop the theory of Dirichlet processes.

For our next result, we will need a couple of facts about the symmetric α-stable proces L_t. It is well known that

$$L_t = \int_0^t \int_{|r|\leq 1} r\widetilde{N}(ds, dr) + \int_0^t \int_{|r|>1} rN(ds, dr). \tag{2.6}$$

Here, $N(ds, dr)$ is the Poisson random measure associated with L and $\widetilde{N}(ds, dr)$ is the compensated Poisson random measure. The compensator measure of N is given by $c_\alpha |r|^{-1-\alpha}\, ds\, dr$.

PROPOSITION 2.8. *Let $\lambda_0 = \lambda_0(\beta, 2\|b\|_\beta)$ be as in Proposition 2.7 and let $\lambda \geq \lambda_0$. Let $X = (X_t)_{t\in[0,T]}$ be a weak solution to (1.1) in the class $\mathcal{V}((1 + \frac{\beta}{\alpha}) \wedge 1)$ and $u^\lambda_b = u^\lambda_{b,b}$ be the* *unique solution to* (2.4). *Then for any* $t \in [0, T]$,

$$u_b^\lambda(X_t) + X_t = u_b^\lambda(x) + x + \lambda \int_0^t u_b^\lambda(X_s)\, ds \\ + \int_0^t \int_{\mathbb{R}} [u_b^\lambda(X_{s-} + r) - u_b^\lambda(X_{s-})] \widetilde{N}(ds, dr) + L_t. \tag{2.7}$$

As mentioned before, we will call the SDE (2.7) the Zvonkin equation. Note that: first, all the terms in the Zvonkin equation make sense; second, (2.7) does not have any distributional drift term like (1.1); and finally, if λ is very large but still finite, then u_b^λ is very close to zero, and thus the only term with X_t that will not disappear in (2.7) (apart from X_t itself) is $\lambda \int_0^t u_b^\lambda(X_s)\, ds$, which is smooth in t and behaves "nicely." To establish (2.7), the Krylov-type estimate is used (see Lemma 5.3).

To show existence of a weak solution for (1.1) and (2.7), we construct an approximating sequence. Let $(b_n)_{n \in \mathbb{Z}_+}$ be a sequence of functions in $\mathcal{C}_b^\infty$ such that b_n converges to b in $\mathcal{C}^{\beta-}$ and $\|b_n\|_\beta \leq 2\|b\|_\beta$. Let $X^n = (X_t^n)_{t \in [0,T]}$, $n \in \mathbb{Z}_+$ be the strong solution to the following stochastic differential equation:

$$X_t^n = x + \int_0^t b_n(X_s^n)\, ds + L_t, \quad t \in [0, T] \tag{2.8}$$

and put $A^n = (A_t^n)_{t \in [0,T]}$, $n \in \mathbb{Z}_+$,

$$A_t^n = \int_0^t b_n(X_s^n)\, ds, \quad t \in [0, T]. \tag{2.9}$$

The strong existence and uniqueness for (2.8) is well known (see, e.g., [2], Theorem 6.2.3). To show tightness and subsequential limit of the above sequence, we will use the Zvonkin transformation. We obtain the following result. Let the Skorokhod space $\mathbb{D}_{\mathbb{R}}[0, T]$ be the space of all càdlàg functions from $[0, T]$ to $\mathbb{R}$.

PROPOSITION 2.9. *Let* $\lambda_0 = \lambda_0(\beta, 2\|b\|_\beta)$ *be as in Proposition* 2.7 *and* $\lambda \geq \lambda_0$. *Let* $u_b^\lambda = u_{b,b}^\lambda$ *be the unique solution to* (2.4) *and let* (X^n, A^n) *be defined as above. Then there exists a subsequence* n_k *such that* (X^{n_k}, A^{n_k}) *converges weakly to* (X, A) *in* $\mathbb{D}_{\mathbb{R}}[0, T]$. *Further*:

(i) X *is a weak solution of the Zvonkin equation* (2.7).
(ii) X *is a weak solution to stochastic differential equation* (1.1) *and it is in the class* $\mathcal{V}((1 + \frac{\beta}{\alpha}) \wedge 1)$.

Our final ingredient is to establish pathwise uniqueness for (2.7).

PROPOSITION 2.10. *There exists* $\lambda_1 = \lambda_1(\beta, \|b\|_\beta)$ *such that for any* $\lambda > \lambda_1$ *the Zvonkin equation* (2.7) *has a pathwise unique solution*.

Strong existence and uniqueness for equations of the type (2.7) with non-Lipschitz coefficients have been studied before, see, for example, [35, 36]. In our proof, we will show that equations of the type (2.7) satisfy the hypothesis of [36], Theorem 3.2, yielding pathwise uniqueness. One could also prove the result directly. For the sake of completeness, we provide such a proof in the Supplementary Material [3], Section 8.

REMARK 2.11. It should be possible to extend the results from Proposition 2.9 to any dimension. However, our argument for proving Proposition 2.10 works only in $d = 1$.

We now have all the key ingredients to complete the proof of Theorem 2.3. Proposition 2.9(ii) shows that (1.1) has a weak solution. Proposition 2.10 and Proposition 2.8 together show that (1.1) has pathwise uniqueness. We then use [31], Theorem 3.4, to establish the classical Yamada–Watanabe theorem, which implies existence and uniqueness of a strong solution in our general setting. We present the details in Section 5.4.

The rest of the paper is organized as follows. In Section 3, we present a number of preliminary results that are used for the proof of the theorem. Most of them are very well known and are provided for the sake of completeness. We discuss the basic properties of the Besov norms, convergence in the Skorokhod space and Dirichlet processes. Section 4 is devoted to the proof of Proposition 2.7. In Section 5, we give the proofs of Propositions 2.8, 2.9, 2.10. This allows to complete the proof of Theorem 2.3 in Section 5.4. Some auxiliary results are proved in the Supplementary Material [3].

Convention on constants. Throughout the paper, C denotes a positive constant whose value may change from line to line. All other constants will be denoted by $C_1, C_2, \dots$ They are all positive and their precise values are not important. The dependence of constants on parameters if needed will be mentioned inside brackets, for example, $C(\alpha, \beta)$.

3. Preliminaries.

3.1. *Besov norms and fractional Laplacian*. In this section, we collect some standard properties of Besov norms and the fractional Laplacian that will be used throughout the paper.

First, we recall that for any $\gamma \in \mathbb{R}$ the space $\mathcal{C}^\gamma$ includes all distributions which are distributional derivatives of elements of $\mathcal{C}^{\gamma+1}$.

The next lemma provides useful properties of Besov norms.

LEMMA 3.1. *Let f be a function $\mathbb{R} \to \mathbb{R}$. Then*:

(i) *For any $\eta, \gamma \in \mathbb{R}$ and $\eta < \gamma$, we have $\|f\|_\eta \le \|f\|_\gamma$.*

(ii) *For any $\eta, \gamma \in \mathbb{R}$ and $\eta < 0 < \gamma$, there exist constants $C_1 > 0$, $C_2 > 0$ such that*

$$\|f\|_\eta \le C_1 \|f\| \le C_2 \|f\|_\gamma.$$

(iii) *For any $\eta \in \mathbb{R}$, there exists $C > 0$ such that*

$$\|f'\|_{\eta-1} \le C\|f\|_\eta. \tag{3.1}$$

(iv) *For any $\eta, \gamma \in \mathbb{R}$, $\eta + \gamma > 0$ there exists $C > 0$ such that for any $g \in \mathcal{C}^\gamma$,*

$$\|fg\|_{\eta\wedge\gamma} \le C\|f\|_\eta \|g\|_\gamma. \tag{3.2}$$

Further, the constants C, C_1, C_2 do not depend on the functions f, g.

PROOF. (i) and (ii) follow, for example, from [37], Exercise 2. (iii) This follows from, for example, [43], Formula 2.3.8.(6). (iv) follows immediately from, for example, [37], Section 2.3 and [37], Theorem 13. □

The properties of Besov norms established in Lemma 3.1 are basic and we will be using them further in the paper without explicit reference to the lemma. Our last lemma in this section describes additional properties of Besov norms in relation to the fractional Laplacian and its associated semigroup.

LEMMA 3.2.

(i) *For any $\gamma \in \mathbb{R}$, there exists $C > 0$ such that for any $f \in \mathcal{C}^\gamma$,*

$$\|\mathcal{L}_\alpha f\|_{\gamma-\alpha} \le C\|f\|_\gamma.$$

(ii) *For any* $\gamma \geq 0$, $\eta \in (-\infty, \gamma]$ *there exists* $C > 0$ *such that for any* $f \in \mathcal{C}_b^\infty$, $t \in (0, 1]$,

$$\|P_t f\|_\gamma \leq C t^{\frac{\eta-\gamma}{\alpha}} \|f\|_\eta$$

(iii) *For any* $\gamma \geq 0$, $\eta \in (-\infty, \gamma]$ *there exists* $C > 0$ *such that for any* $f \in \mathcal{C}_b^\infty$, $t \geq 1$,

$$\|P_t f\|_\gamma \leq C \|f\|_\eta.$$

(iv) *For any* $f \in \mathcal{C}_b^\infty$,

$$\int_{\mathbb{R}} |\mathcal{L}_\alpha f(x)| \, dx \leq 4c_\alpha \int_{\mathbb{R}} |f(x)| \, dx + \frac{2c_\alpha}{2-\alpha} \int_{\mathbb{R}} \Bigl(\sup_{z \in [x-1, x+1]} |f''(z)| \Bigr) \, dx,$$

where the constant c_α *was defined in* (2.3).

PROOF. (i), (iii), (iv). These statements are standard however we were not able to find their proofs in the literature; for the sake of completeness, we provide their proofs in the Supplementary Material [3], Section 2. (ii) We refer the reader to [22]. Even though the statement of Lemma A.7 is for a bounded set, one can verify that the proof works also for $\mathbb{R}$. $\square$

3.2. *Properties of convergence in the Skorokhod space*. Let E be a metric space. In this section, we provide some technical though important tools for studying convergence in the Skorokhod space $\mathbb{D}_E[0, T]$, that is, the space of all càdlàg functions from $[0, T]$ to E. We refer the reader to [8], Chapter 3 and [16], Chapter 3 for a detailed treatment of the Skorokhod space and the necessary definitions. Denote by d the Skorokhod distance in $\mathbb{D}_E$. Let Λ be the set of continuous strictly increasing functions mapping $[0, T]$ onto $[0, T]$.

We will use the following lemma, which is just a minor modification of a corresponding lemma from [32]. For the convenience of the reader and for the sake of exposition, we state this proposition and its proof below.

LEMMA 3.3 (cf. [32], Lemma 2.1). *Let* E_1 *and* E_2 *be metric spaces and let* Φ, Φ_n, *where* $n \in \mathbb{Z}_+$, *be mappings* $\mathbb{D}_{E_1}[0, T] \to \mathbb{D}_{E_2}[0, T]$, $T > 0$. *Suppose that for any* $n \in \mathbb{Z}_+$ *we have* $\Phi^n(Z \circ \mu) = \Phi^n(Z) \circ \mu$ *whenever* $Z \in \mathbb{D}_{E_1}[0, T]$, $\mu \in \Lambda$. *Further assume that* $Z^n \to Z$ *in the uniform metric implies* $\Phi^n(Z^n) \to \Phi(Z)$ *in the uniform metric. Then* $Z^n \to Z$ *in the Skorokhod topology implies* $(Z^n, \Phi^n(Z^n)) \to (Z, \Phi(Z))$ *in the Skorokhod topology.*

PROOF. The proof follows the proof of [32], Lemma 2.1, with minor modifications. Let ρ_{E_1} and ρ_{E_2} denote the metric on E_1 and E_2, respectively. Let $\rho_{E_1 \times E_2}$ be the product metric on $E_1 \times E_2$.

Take any $Z \in \mathbb{D}_{E_1}[0, T]$, a sequence $(Z^n)_{n \in \mathbb{Z}_+}$, $Z^n \in \mathbb{D}_{E_1}[0, T]$ and assume that $d(Z^n, Z) \to 0$. Then there exists a mapping $\mu^n \in \Lambda$ such that $\|\mu^n - I\| \to 0$ where I denote the identity map and $\sup_{t \in [0,T]} \rho_{E_1}((Z^n \circ \mu^n)(t), Z_t) \to 0$. By the assumptions, this implies that

$$\sup_{t \in [0,T]} \rho_{E_2}\bigl(\Phi^n\bigl(Z^n \circ \mu^n\bigr)(t), \Phi(Z)(t)\bigr) \to 0.$$

Since $\Phi^n(Z^n \circ \mu^n) = \Phi^n(Z^n) \circ \mu^n$, we finally obtain

$$\begin{aligned}
&d\bigl(\bigl(Z^n, \Phi^n\bigl(Z^n\bigr)\bigr), \bigl(Z, \Phi(Z)\bigr)\bigr) \\
&\quad \leq \|\mu^n - I\| + \sup_{t \in [0,T]} \rho_{E_1 \times E_2}\bigl(\bigl(\bigl[Z^n \circ \mu^n\bigr](t), \bigl[\Phi^n\bigl(Z^n\bigr) \circ \mu^n\bigr](t)\bigr), \bigl(Z_t, \Phi(Z)(t)\bigr)\bigr) \\
&\quad \to 0.
\end{aligned}$$

This implies the statement of the lemma. $\square$

We will use the following simple lemma that deals with convergence of integrals in the Skorokhod space.

LEMMA 3.4. *Let $(X^n)_{n\in\mathbb{Z}_+}$ be a sequence of elements in $\mathbb{D}_{\mathbb{R}}[0,T]$ converging a.s. in the Skorokhod metric to X. Let $(f_n)_{n\in\mathbb{Z}_+}$ be a sequence of continuous functions $\mathbb{R}^2\to\mathbb{R}$ converging uniformly to f. Assume*

$$\sup_{n\in\mathbb{Z}_+}\|f_n\|<\infty.$$

Let $A\in\mathbb{R}$ be a Borel measurable set and let θ be a finite measure on A.

Then the process J^n defined as

$$J^n(t):=\int_0^t\int_A f_n(X^n_{s-},r)\theta(dr)\,ds,\quad t\in[0,T],$$

uniformly on $[0,T]$ converges to the process J

$$J(t):=\int_0^t\int_A f(X_{s-},r)\theta(dr)\,ds,\quad t\in[0,T]$$

a.s. as $n\to\infty$.

PROOF. We have

$$\|J^n-J\|\le\int_0^T\int_A\big|f_n(X^n_{s-},r)-f(X_{s-},r)\big|\theta(dr)\,ds.$$

Recall that X^n converges to X a.s. in the Skorokhod metric. Therefore, for almost all $\omega\in\Omega$ the sequence $(X^n_{t-}(\omega))_{n\in\mathbb{Z}_+}$ converges to $X_{t-}(\omega)$ for all but countably many $t\in[0,T]$. Since the sequence $(f_n)_{n\in\mathbb{Z}_+}$ is uniformly bounded and converges pointwise to f, we can apply the dominated convergence theorem to conclude

$$\|J^n-J\|\to 0\quad\text{a.s. as } n\to\infty.\qquad\square$$

The next lemma is more complicated and it deals with passing to the limit in stochastic integrals.

LEMMA 3.5. *Let $(X^n,L^n)_{n\in\mathbb{Z}_+}$ be a sequence of elements in $\mathbb{D}_{\mathbb{R}^2}[0,T]$ converging a.s. in the Skorokhod metric to (X,L). Let ν be a Levy measure satisfying*

$$\int_{\mathbb{R}}(1\wedge r^2)\nu(dr)<\infty. \tag{3.3}$$

Assume that L, $(L^n)_{n\in\mathbb{Z}_+}$ are Lévy processes with the Lévy measure ν. Let N^n and $\widetilde{N}^n$ be the Poisson measure and compensated Poisson measure associated with L^n, respectively, where $n\in\mathbb{Z}_+$. Define N and $\widetilde{N}$ in a similar way. Let $(\mathcal{F}^{X^n}_t)_{t\ge0}$ (resp., $(\mathcal{F}^X_t)_{t\ge0}$) be the natural filtration of X^n (resp., X). Assume that for any compact set $A\subset\mathbb{R}\setminus\{0\}$ and $0\le s<t\le T$ the random variable $\widetilde{N}^n_t(A)-\widetilde{N}^n_s(A)$ is independent of $\mathcal{F}^{X^n}_s\vee\mathcal{F}^{\widetilde{N}^n}_s$. Then:

(i) *For any compact set $A\subset\mathbb{R}\setminus\{0\}$ and $0\le s<t\le T$, the random variable $\widetilde{N}_t(A)-\widetilde{N}_s(A)$ is independent of $\mathcal{F}^X_s\vee\mathcal{F}^{\widetilde{N}}_s$.*

(ii) *Let $(f_n)_{n\in\mathbb{Z}_+}$ be a sequence of continuous functions $\mathbb{R}^2\to\mathbb{R}$ converging uniformly to f. Assume that for some $C>0$,*

$$\big|f_n(x,r)\big|\le C\big(|r|\wedge 1\big),\quad x,r\in\mathbb{R},n\in\mathbb{Z}_+. \tag{3.4}$$

Then for any $T>0$ *the process* I^n *defined as*

$$I^n(t) := \int_0^t \int_{\mathbb{R}} f_n(X^n_{s-}, r)\widetilde{N}^n(ds, dr), \quad t \in [0, T]$$

converges in probability in the Skorokhod space $\mathbb{D}_{\mathbb{R}}[0, T]$ *to the process* I,

$$I(t) := \int_0^t \int_{\mathbb{R}} f(X_{s-}, r)\widetilde{N}(ds, dr), \quad t \in [0, T].$$

The proof of this lemma is provided in the Supplementary Material [3], Section 3.

3.3. *Dirichlet processes.* Recall that a weak or strong solution to SDE (1.1) is not necessarily a semimartingale; it belongs to a more general class of processes called Dirichlet processes. Thus to study properties of solutions of our equation we need to develop some parts of the theory of Dirichlet processes. This is done in this section. We begin with the following definitions.

DEFINITION 3.6 ([21]). We say that a continuous adapted process $(A_t)_{t\in[0,T]}$ is a *process of zero energy* if $A_0 = 0$ and

$$\lim_{\delta\to 0} \sup_{\pi_T : |\pi_T|<\delta} \mathsf{E}\Big(\sum_{t_i\in\pi_T} |A_{t_{i+1}} - A_{t_i}|^2\Big) = 0,$$

where π_T denotes a finite partition of $[0, T]$ and $|\pi_T|$ denotes the mesh size of the partition.

DEFINITION 3.7 ([21]). We say that an adapted process $(X_t)_{t\in[0,T]}$ is a *Dirichlet process* if

$$X_t = M_t + A_t, \quad t \in [0, T], \tag{3.5}$$

where M is a square–integrable martingale and A is an adapted process of zero energy.

It was proven in [21] that such decomposition (3.5) of a Dirichlet process X is unique. Thus we see that the class of Dirichlet processes naturally extends the class of semimartingales. Note however that since the process A in decomposition (3.5) might be of infinite variation, the integral with respect to A might be not well defined in the classical sense. The next definition extends the notion of a stochastic integral to the class of integrals with respect to zero-energy processes. We shall define the integral with respect to the zero-energy process A as a limit in probability of the corresponding forward Riemann sums.

DEFINITION 3.8 ([14], page 90). Let $f : \mathbb{R} \to \mathbb{R}$ be a bounded continuous function with a bounded continuous derivative. Let X be a Dirichlet process with decomposition (3.5). For $n \in \mathbb{Z}_+$, let $D_n := \{t_i^n\}$ be a sequence of *refining* (and nonrandom) partitions of $[0, T]$ whose mesh size tends to 0 as $n \to \infty$. Then

$$\int_s^t f(X_r)\, dA_r := (\mathsf{P}) \lim_{n\to\infty} \sum_{t_i^n\in D_n, t_i^n\in[s,t)} f(X_{t_i^n})(A_{t_{i+1}^n} - A_{t_i^n}), \quad 0 \le s \le t \le T$$

(if exists), where the limit is taken in probability.

It was shown in [14], Theorem 3.1, that if f is a bounded continuous function with a bounded continuous derivative, then the integral $\int_0^\cdot f(X_s)\, dA_s$ exists and does not depend on the choice of the sequence of refining partitions D_n.

We will need a couple of statements describing further properties of integrals with respect to Dirichlet processes. Some of the results are close in spirit to [6], Lemma 2.3 and [45], Lemma 3.12 and thus their proofs are moved to the Supplementary Material [3].

In the first lemma, we prove that under certain regularity conditions, the convergence in probability in the definition of the Dirichlet integral can be improved to convergence in $\mathbb{L}_p(\Omega)$.

LEMMA 3.9. *Let (X, A) be as in Definition 3.7. Let $f : \mathbb{R} \to \mathbb{R}$ be a bounded continuous function with a bounded continuous derivative. Suppose that for some $p_1, p_2 > 0$ and $\gamma_1, \gamma_2 > 0$ with $\gamma_1 + \gamma_2 > 1$ and $1/h := 1/p_1 + 1/p_2 \le 1$ there exist constants $C_f, C_A \ge 1$ such that for any $s, t \in [0, T]$*

$$\mathsf{E}\big|f(X_t) - f(X_s)\big|^{p_1} \le (C_f)^{p_1}|t-s|^{p_1\gamma_1}, \qquad \mathsf{E}\big|f(X_t)\big|^{p_1} \le (C_f)^{p_1}, \tag{3.6}$$

$$\mathsf{E}|A_t - A_s|^{p_2} \le (C_A)^{p_2}|t-s|^{p_2\gamma_2}. \tag{3.7}$$

(i) *Then for any $0 \le s \le t \le T$ the sequence of partial sums*

$$I_n := \sum_{i=0}^{2^n-1} f(X_{t_n^i})(A_{t_n^{i+1}} - A_{t_n^i}),$$

where

$$t_n^k := s + k2^{-n}(t-s) \quad \text{for } k = 0, 1, \dots, 2^n; n \in \mathbb{Z}_+,$$

converges to $I := \int_s^t f(X_r)\, dA_r$ in $\mathbb{L}_h = \mathbb{L}_h(\Omega)$.

(ii) *Moreover, there exists a constant $C = C(T, \gamma_1, \gamma_2) > 0$ such that for any $0 \le s \le t \le T$, $n \in \mathbb{Z}_+$, we have the following estimate of the remainder term*:

$$\|I - I_n\|_{\mathbb{L}_h} \le CC_f C_A 2^{-n(\gamma_1+\gamma_2-1)}. \tag{3.8}$$

(iii) *Finally, there exists $C = C(T, \gamma_1, \gamma_2) > 0$ such that for any $0 \le s \le t \le T$,*

$$\left\| \int_s^t f(X_r)\, dA_r \right\|_{\mathbb{L}_h} \le CC_f C_A (t-s)^{\gamma_2}. \tag{3.9}$$

PROOF. We present the proof in the Supplementary Material [3], Section 4. □

The second lemma of this subsection deals with the approximations of the integral with respect to a Dirichlet process.

LEMMA 3.10. *Let (X, A) be as in Definition 3.7. Let $(f_n)_{n\in\mathbb{Z}_+}$ be a sequence of functions $\mathbb{R} \to \mathbb{R}$ that are uniformly bounded, continuous and have a bounded continuous derivative. Assume that all f_n satisfy condition (3.6) with the same parameters p_1, γ_1 and C_{f_1} for all $n \in \mathbb{Z}_+$.*

Let $(b_n)_{n\in\mathbb{Z}_+}$ be a sequence of bounded continuous functions. Define

$$\mathcal{A}_t^n := \int_0^t b_n(X_s)\, ds, \quad t \in [0, T].$$

Suppose that for each $t \in [0, T]$ the sequence $(\mathcal{A}^n(t))_{n\in\mathbb{Z}_+}$ converges in probability to $A(t)$.

Assume that A and all functions $\mathcal{A}^n$, $n \in \mathbb{Z}_+$ satisfy condition (3.7) with the same parameters C_A, p_2, γ_2.

Finally, assume that $\gamma_1 + \gamma_2 > 1$ and $1/p_1 + 1/p_2 \le 1$.

Then for any $t \in [0, T]$ we have

$$\int_0^t f_n(X_s)\, dA_s - \int_0^t f_n(X_s) b_n(X_s)\, ds \to 0 \quad \text{in probability as } n \to \infty. \tag{3.10}$$

PROOF. We present the proof in the Supplementary Material [3], Section 5. □

4. Proof of Proposition 2.7: Analysis of the resolvent equation. The primary purpose of this section is to prove Proposition 2.7. We will follow an approach similar to [40]. We begin the analysis of equation (2.4) with the case $f=0$.

LEMMA 4.1. *The resolvent equation* (2.4) *with* $\lambda \geq 1$, $f=0$, $g \in \mathcal{C}^\eta$, $\eta > -\alpha$ *has a unique solution in the class of bounded functions. Furthermore, for each* $\gamma \in [0 \vee \eta, \alpha+\eta)$, *this solution* $u^\lambda_{0,g} \in \mathcal{C}^\gamma$ *and there exists a constant* $C = C(\eta,\gamma) > 0$ *such that*

$$\|u^\lambda_{0,g}\|_\gamma \leq C\lambda^{-1-\eta/\alpha+\gamma/\alpha}\|g\|_\eta. \tag{4.1}$$

PROOF. We begin with uniqueness. Let u_1, u_2 be two bounded solutions of (2.4) with $\lambda > 0$, $f=0$, $g \in \mathcal{C}^\eta$. Then the function $v := u_1 - u_2$ is obviously bounded and we have $\mathcal{L}_\alpha v = \lambda v$. Take any test function $\phi \in \mathcal{S}$. It follows from the definition of the solution that

$$\langle v, \mathcal{L}_\alpha \phi - \lambda\phi\rangle = 0. \tag{4.2}$$

We claim now that for any $h \in \mathcal{C}_c^\infty$ one has

$$\langle v, h\rangle = 0.$$

To prove it, we fix any $h \in \mathcal{C}_c^\infty$ and put

$$\psi(x) := \int_0^\infty e^{-\lambda t} P_t h(x)\,dt = \int_{\mathbb{R}}\int_0^\infty e^{-\lambda t} p_t(x-y)\,dt\,h(y)\,dy, \quad x \in \mathbb{R}, \tag{4.3}$$

where p_t is the α-stable transition density.

Now let us use the standard estimates on α-stable transition density (see, e.g., [9], Theorem 2.1) and its self-similarity, $p_t(x) = t^{-1/\alpha}p_1(t^{-1/\alpha}x)$, to get

$$p_t(x) \leq c\min\{t^{-1/\alpha}, t|x|^{-1-\alpha}\} \quad \text{for any } x \in \mathbb{R},$$

for some constant $c > 0$. Therefore, we easily derive for any $x \in \mathbb{R}$,

$$\begin{aligned}\int_0^\infty e^{-\lambda t} p_t(x)\,dt &\leq C\int_0^\infty e^{-\lambda t}\min\{t^{-1/\alpha}, t|x|^{-1-\alpha}\}\,dt \\ &\leq C(\lambda)\min\{1, |x|^{-1-\alpha}\},\end{aligned} \tag{4.4}$$

for some constant $C(\lambda) > 0$. For $p > 0$, define the following normed space of functions on $\mathbb{R}$:

$$\mathcal{C}_{\mathrm{rap},p} \equiv \Big\{f \in \mathcal{C}_b : \|f\|_{\mathrm{rap},p} := \sup_{x\in\mathbb{R}}|f(x)|\max\{1, |x|^p\} < \infty\Big\},$$

where $\mathcal{C}_b$ is the set of bounded continuous functions on $\mathbb{R}$. From (4.3) and (4.4), one can immediately derive that

$$\psi, \psi', \psi'' \in \mathcal{C}_{\mathrm{rap},1+\alpha}.$$

Moreover, one can check that $\psi, \psi', \psi'' \in \mathbb{L}_1(\mathbb{R}) \cap \mathcal{C}_b^\infty$ and also (see, e.g., [40], Theorem 3.3) that

$$\mathcal{L}_\alpha\psi - \lambda\psi = h.$$

Fix arbitrary $\varepsilon \in (0,1)$. Let now ψ_n be a sequence of $\mathcal{C}_c^\infty$ functions, such that ψ_n, ψ_n', ψ_n'' converge to ψ, ψ', ψ'' in $\mathbb{L}_1(\mathbb{R})$ and in $\mathcal{C}_{\mathrm{rap},1+\alpha-\varepsilon}$, respectively. Then we get

$$\int_{\mathbb{R}}\Big(\sup_{z\in[x-1,x+1]}|\psi_n''(z) - \psi''(z)|\Big)\,dx \to 0.$$

Hence by Lemma 3.2(iv) we have

$$\int_{\mathbb{R}} \left|\mathcal{L}_\alpha \psi_n(x) - \mathcal{L}_\alpha \psi(x)\right| dx \to 0.$$

Therefore, using (4.2) and the fact that the function v is bounded, we get

$$\begin{aligned}\left|\langle v, h\rangle\right| &= \left|\langle v, \mathcal{L}_\alpha \psi - \lambda\psi\rangle\right| \\ &\le \left|\langle v, \mathcal{L}_\alpha \psi_n - \lambda\psi_n\rangle\right| + \left|\langle v, \lambda\psi_n - \lambda\psi\rangle\right| + \left|\langle v, \mathcal{L}_\alpha \psi_n - \mathcal{L}_\alpha\psi\rangle\right| \to 0.\end{aligned}$$

Thus

$$\langle v, h\rangle = 0$$

for all $h \in \mathcal{C}_c^\infty$. This yields $v = 0$ and completes the proof of uniqueness.

To show existence of solution and to establish estimate (4.1), we adapt some ideas from the proof of [40], Theorem 3.3. Fix $\gamma \in [0 \vee \eta, \alpha + \eta)$ and take any $\alpha' \in (\alpha, 2)$. We begin with the case $g \in \mathcal{C}_b^\infty$. It was shown in [40], Theorem 3.3 that in this case equation (2.4) with $f = 0$ has a unique solution in $\mathcal{C}^{\alpha'}$. This solution is given by

$$u_{0,g}^\lambda(x) := \int_0^\infty e^{-\lambda t} P_t g(x)\, dt, \quad x \in \mathbb{R},$$

where the semigroup $(P_t)_{t\ge 0}$ is as in Definition 2.6. Hence, using Lemma 3.2(ii), (iii), we obtain

$$\begin{aligned}\|u_{0,g}^\lambda\|_\gamma &\le \int_0^{+\infty} e^{-\lambda t} \|P_t g\|_\gamma\, dt \\ &= \int_0^1 e^{-\lambda t}\|P_t g\|_\gamma\, dt + \int_1^{+\infty} e^{-\lambda t}\|P_t g\|_\gamma\, dt \\ &\le C\|g\|_\eta \int_0^1 e^{-\lambda t} t^{-(\gamma-\eta)/\alpha}\, dt + C\|g\|_\eta \int_1^\infty e^{-\lambda t}\, dt \\ &\le C\lambda^{-1-\eta/\alpha+\gamma/\alpha}\|g\|_\eta,\end{aligned} \tag{4.5}$$

where the last inequality follows from the fact that $\lambda \ge 1$.

Now take any $g \in \mathcal{C}^\eta$ and fix arbitrary $\eta' \in (\gamma - \alpha, \eta)$. Let $g_n \in \mathcal{C}_b^\infty$, $n \in \mathbb{Z}_+$ be a sequence of approximations of g such that $\|g_n - g\|_{\eta'} \to 0$ as $n \to \infty$ and

$$\sup_n \|g_n\|_\eta \le 2\|g\|_\eta. \tag{4.6}$$

Consider the function $u_n := u_{0,g_n}^\lambda$. By above, u_n is well defined and $u_n \in \mathcal{C}^{\alpha'}$.

Let $v_{n,m} := u_n - u_m$, $n, m \in \mathbb{Z}_+$. We see that $v_{n,m}$ solves (2.4) with $f = 0$ and the right-hand side $g_n - g_m$. Furthermore, since $u_n, u_m \in \mathcal{C}^{\alpha'}$, we see that $v_{n,m} \in \mathcal{C}^{\alpha'}$. Recall that the solution to (2.4) with $f = 0$ and smooth right-hand side is unique in class $\mathcal{C}^{\alpha'}$ by [40], Theorem 3.3. Therefore, we can apply (4.5) (with η' instead of η) to obtain

$$\|u_n - u_m\|_\gamma = \|v_{n,m}\|_\gamma = \left\|u_{0,g_n-g_m}^\lambda\right\|_\gamma \le C\lambda^{-1-\eta'/\alpha+\gamma/\alpha}\|g_n - g_m\|_{\eta'}.$$

This implies that $(u_n)_{n\in\mathbb{Z}_+}$ is a Cauchy sequence in $\mathcal{C}^\gamma$, and hence there exists some $u \in \mathcal{C}^\gamma$ such that $\|u - u_n\|_\gamma \to 0$ as $n \to \infty$. We claim that u is a solution to (2.4) with $f = 0$ and the right-hand side g. Indeed,

$$\begin{aligned}\|\lambda u - \mathcal{L}_\alpha u - g\|_{\gamma-\alpha} &= \left\|\lambda(u - u_n) - \mathcal{L}_\alpha(u - u_n) - (g - g_n)\right\|_{\gamma-\alpha} \\ &\le \lambda\|u - u_n\|_{\gamma-\alpha} + \left\|\mathcal{L}_\alpha(u - u_n)\right\|_{\gamma-\alpha} + \|g - g_n\|_{\gamma-\alpha} \\ &\le \lambda\|u - u_n\|_\gamma + C\|u - u_n\|_\gamma + \|g - g_n\|_{\eta'},\end{aligned}$$

where we used the fact that $\gamma - \alpha < \eta'$, Lemma 3.1, and Lemma 3.2. By passing to the limit as $n \to \infty$, we deduce

$$\lambda u - \mathcal{L}_\alpha u - g = 0,$$

and hence u indeed solves (2.4) with $f = 0$ and the right-hand side g. To complete the proof, it remains to note that by (4.5),

$$\|u\|_\gamma \le \|u_n\|_\gamma + \|u - u_n\|_\gamma \le C\lambda^{-1-\eta/\alpha+\gamma/\alpha}\|g_n\|_\eta + \|u - u_n\|_\gamma.$$

By passing to the limit as $n \to \infty$ and using (4.6), we obtain (4.1). □

Now we are ready to prove the first part of Proposition 2.7.

PROOF OF PROPOSITION 2.7(i). We begin by proving a crucial inequality that will be used many times in this proposition. Let $\delta > -\alpha$ and let u be any bounded solution to (2.4) with $f, g \in \mathcal{C}^\delta$, $\lambda \ge 1$. Obviously,

$$\lambda u - \mathcal{L}_\alpha u = f u' + g. \tag{4.7}$$

Let $\rho \in \mathbb{R}$ be such that $\rho > -\delta$ and $\rho \ge \delta$. Using (4.7) and Lemma 4.1, we derive that for any $\gamma \in [0 \vee \delta, \alpha + \delta)$ there exists $C = C(\delta, \gamma, \rho) > 0$ such that

$$\begin{aligned} \|u\|_\gamma &\le C\lambda^{-1-\frac{\delta}{\alpha}+\frac{\gamma}{\alpha}}\|f u' + g\|_\delta \\ &\le C\lambda^{-1-\frac{\delta}{\alpha}+\frac{\gamma}{\alpha}}\big(\|g\|_\delta + \|f\|_\delta\|u'\|_\rho\big) \\ &\le C\lambda^{-1-\frac{\delta}{\alpha}+\frac{\gamma}{\alpha}}\big(\|g\|_\delta + \|f\|_\delta\|u\|_{\rho+1}\big), \end{aligned} \tag{4.8}$$

whenever $\delta + \rho > 0$ and $\rho \ge \delta$. Here, we have used inequalities from Section 3.1, specifically (3.2) in the second inequality and (3.1) in the third inequality. We will apply (4.8) repeatedly. Now we can start proving (2.5).

First, we deal with the case $\eta > 0$, and take $f, g \in \mathcal{C}^\eta$. By [40], Theorem 3.4, equation (2.4) has a solution $u^\lambda_{f,g} \in \mathcal{C}^\alpha$ for all $\lambda > 0$. We will show that

$$u^\lambda_{f,g} \in \mathcal{C}^\gamma \qquad \text{for any } \lambda > 0, \gamma < \alpha + \eta. \tag{4.9}$$

Assume the converse. Then there exist $\alpha \le \gamma_1 < \gamma_2 < \alpha + \eta$, such that $\|u^\lambda_{f,g}\|_{\gamma_1} < \infty$, $\|u^\lambda_{f,g}\|_{\gamma_2} = \infty$ and $\gamma_2 - \gamma_1 < \alpha - 1$. We apply (4.8) with $\gamma = \gamma_2$, $\delta = (\gamma_1 - 1) \wedge \eta$, $\rho = \gamma_1 - 1$. Note that all the additional constraints are satisfied: since $\gamma_2 < \alpha + \eta$ and $\gamma_2 < \alpha + \gamma_1 - 1$, we see that $\gamma_2 < \alpha + (\gamma_1 - 1) \wedge \eta$, and thus $\gamma_2 \in [0 \vee \delta, \alpha + \delta)$. Using the fact that by assumption $f, g \in \mathcal{C}^\eta$, we derive

$$\|u^\lambda_{f,g}\|_{\gamma_2} \le C(\lambda)\big(\|g\|_{(\gamma_1-1)\wedge\eta} + \|f\|_{(\gamma_1-1)\wedge\eta}\|u^\lambda_{f,g}\|_{\gamma_1}\big) < \infty,$$

However, this contradicts the assumption $\|u^\lambda_{f,g}\|_{\gamma_2} = \infty$. Thus $u^\lambda_{f,g} \in \mathcal{C}^\gamma$ for any $\gamma < \alpha + \eta$.

We apply (4.8) again, but now with a different set of parameters: we take $\gamma = \eta + 1$, $\delta = \rho = \eta$. Then we obtain

$$\|u^\lambda_{f,g}\|_{\eta+1} \le C_1\lambda^{-1+\frac{1}{\alpha}}\big(\|g\|_\eta + \|f\|_\eta\|u^\lambda_{f,g}\|_{\eta+1}\big), \tag{4.10}$$

where $C_1 > 0$. By above, $\|u^\lambda_{f,g}\|_{\eta+1} < \infty$. Take now $\lambda_1 \ge 1$ such that

$$1 - C_1\lambda_1^{-1+\frac{1}{\alpha}}M \ge 1/2.$$

Since C_1 depends only on η, we see that λ_1 depends only on η, M. For $\lambda \ge \lambda_1$ and $\|f\|_\eta \le M$, we get from (4.10),

$$\|u^\lambda_{f,g}\|_{\eta+1} \le 2C_1\lambda^{-1+\frac{1}{\alpha}}\|g\|_\eta \le 2C_1\|g\|_\eta. \tag{4.11}$$

Finally, applying again (4.8) with $\delta = \rho = \eta$ and $\gamma \in [\eta, \alpha+\eta)$ we get for $\lambda \ge \lambda_1$

$$\|u^\lambda_{f,g}\|_\gamma \le C\lambda^{-1-\frac{\eta}{\alpha}+\frac{\gamma}{\alpha}}\|g\|_\eta\big(\|f\|_\eta + 1\big),$$

where we used bound (4.11). This establishes (2.5) for $\eta > 0$ when $\lambda \ge \lambda_1$.

Now we can treat the case $\eta \le 0$. The problem here is that we do not know a priori that in this case (2.4) has a solution. Therefore, we have to study approximations. We will still use (4.8) as a main tool however with a different set of parameters (since η is negative we cannot take $\delta = \rho = \eta$).

Thus we start with considering any $f, g \in \mathcal{C}^\infty_b$, $\lambda > 0$. By (4.9), equation (2.4) has a solution $u^\lambda_{f,g} \in \mathcal{C}^\infty_b$. We apply (4.8) with $\gamma = \alpha/2 + 1/2$, $\delta = \eta$, and $\rho = \alpha/2 - 1/2$. Since $\eta > 1/2 - \alpha/2$, one can easily see that all the additional constraints on the parameters in (2.4) are satisfied. We get

$$\|u^\lambda_{f,g}\|_{(1+\alpha)/2}\big(1 - C_2\lambda^{-\frac{1}{2}+\frac{1}{2\alpha}-\frac{\eta}{\alpha}}\|f\|_\eta\big) \le C_2\lambda^{-\frac{1}{2}+\frac{1}{2\alpha}-\frac{\eta}{\alpha}}\|g\|_\eta,$$

where $C_2 > 0$ and we have used the fact that $\|u^\lambda_{f,g}\|_{(1+\alpha)/2} < \infty$. Choose $\lambda_2 \ge 1$ such that

$$1 - C_2\lambda_2^{-\frac{1}{2}+\frac{1}{2\alpha}-\frac{\eta}{\alpha}}M \ge 1/2.$$

Similarly, λ_2 depends only η, M. For $\lambda \ge \lambda_2 = \lambda_2(\eta, M)$ and $\|f\|_\eta \le M$ we get

$$\|u^\lambda_{f,g}\|_{(1+\alpha)/2} \le 2C_2\|g\|_\eta. \tag{4.12}$$

Now we take any $f, g \in \mathcal{C}^\eta$. Similar to the proof of Lemma 4.1, we fix arbitrary $\eta' \in (\frac{1-\alpha}{2}, \eta)$ and approximate f and g by the sequences $f_n, g_n \in \mathcal{C}^\infty_b$, correspondingly, such that

$$\|f_n - f\|_{\eta'} \to 0, \qquad \|g_n - g\|_{\eta'} \to 0 \quad \text{as } n \to \infty,$$
$$\|f_n\|_\eta \le 2\|f\|_\eta, \qquad \|g_n\|_\eta \le 2\|g\|_\eta \quad \text{for all } n \in \mathbb{Z}_+.$$

Consider the function $u_n := u^\lambda_{f_n,g_n}$. By above, $u_n \in \mathcal{C}^\infty_b$. Put $v_{n,m} := u_n - u_m$, $n, m \in \mathbb{Z}_+$. It follows that $v_{n,m} \in \mathcal{C}^\infty_b$ and solves

$$\lambda v_{n,m} - \mathcal{L}_\alpha v_{n,m} - f_n v'_{n,m} = g_n - g_m + (f_n - f_m)u'_m. \tag{4.13}$$

Clearly, the right-hand side of (4.13) is in $\mathcal{C}^\infty_b$. Therefore, using the uniqueness theorem for equation (2.4) with smooth coefficients ([40], Theorem 3.4) we see that $v_{n,m}$ is the unique solution to (4.13). Thus we can apply bound (4.12) (with η' instead of η). We make use of the fact that $\|f_n\|_\eta \le 2\|f\|_\eta$ to get for $\lambda \ge \lambda_2(\eta', 2M)$ and $\|f\|_\eta \le M$,

$$\begin{aligned}\|u_n - u_m\|_{(1+\alpha)/2} = \|v_{n,m}\|_{(1+\alpha)/2} &\le C\big(\|g_n - g_m\|_{\eta'} + \big\|(f_m - f_n)u'_m\big\|_{\eta'}\big)\\ &\le C\big(\|g_n - g_m\|_{\eta'} + \|f_m - f_n\|_{\eta'}\|u'_m\|_{\alpha/2-1/2}\big)\\ &\le C\big(\|g_n - g_m\|_{\eta'} + \|f_m - f_n\|_{\eta'}\|g_m\|_{\eta'}\big),\end{aligned}$$

where in the final inequality we used bound (4.12) once again, this time with η. Recalling that $\|g_m\|_\eta \le 2\|g\|_\eta$, we see that the sequence $(u_n)_{n\in\mathbb{Z}_+}$ is a Cauchy sequence in $\mathcal{C}^{(1+\alpha)/2}$. Hence there exists $u \in \mathcal{C}^{(1+\alpha)/2}$ such that $\|u_n - u\|_{(1+\alpha)/2} \to 0$ as $n \to \infty$. Applying Lemma 3.2, we derive for any $n \in \mathbb{Z}_+$,

$$\begin{aligned}
&\|\lambda u - \mathcal{L}_\alpha u - f u' - g\|_{(1-\alpha)/2} \\
&\quad = \|\lambda(u - u_n) - \mathcal{L}_\alpha(u - u_n) - f(u - u_n)' - (g - g_n)\|_{(1-\alpha)/2} \\
&\quad \le \lambda\|u - u_n\|_{(1+\alpha)/2} + C\|u - u_n\|_{(1+\alpha)/2} \\
&\qquad + C\|f\|_\eta \|u - u_n\|_{(1+\alpha)/2} + C\|g - g_n\|_\eta.
\end{aligned}$$

After taking the limit as $n \to \infty$, we see that u solves (2.4).

Note that thanks to (4.12),

$$\|u\|_{(1+\alpha)/2} \le \lim_{n\to\infty} \|u - u_n\|_{(1+\alpha)/2} + \limsup_{n\to\infty} \|u_n\|_{(1+\alpha)/2} \le C\|g\|_\eta.$$

Using this inequality and the fact that u solves (2.4), we apply again (4.8) with $\gamma \in [0, \alpha + \eta)$, $\delta = \eta$, $\rho = \alpha/2 - 1/2$ to obtain

$$\|u\|_\gamma \le C\lambda^{-1-\frac{\eta}{\alpha}+\frac{\gamma}{\alpha}}\big(\|g\|_\eta + C\|f\|_\eta\|g\|_\eta\big).$$

This establishes (2.5) for $\eta \le 0$ when $\lambda > \lambda_2(\eta', 2M)$. Set

$$\lambda_3(\eta, M) := \begin{cases} \lambda_1(\eta, M) & \text{if } \eta > 0, \\ \lambda_2(\eta', 2M) & \text{if } \eta \le 0, \end{cases}$$

to complete the proof of (2.5).

Thus it remains to show uniqueness. If $u_1, u_2 \in \mathcal{C}^{(1+\alpha)/2}$ are two solutions of (2.4), then for $v := u_1 - u_2$ we obviously have $v \in \mathcal{C}^{(1+\alpha)/2}$ and

$$\lambda v - \mathcal{L}_\alpha v = f v'.$$

Therefore, the right-hand side of the above equation is well defined and is in $\mathcal{C}^\eta$. Therefore, we can apply (4.8) with $\gamma = \alpha/2 + 1/2$, $\delta = \eta$, $\rho = \alpha/2 - 1/2$ to obtain for some $C_3 > 0$,

$$\|v\|_{(1+\alpha)/2} \le C_3\lambda^{-\frac{1}{2}+\frac{1}{2\alpha}-\frac{\eta}{\alpha}}\|f\|_\eta\|v\|_{(1+\alpha)/2}.$$

Choose $\lambda_4 \ge 1$ such that

$$C_3\lambda_4^{-\frac{1}{2}+\frac{1}{2\alpha}-\frac{\eta}{\alpha}}\|M\|_\eta \le 1/2.$$

Then, by above, $\|v\|_{(1+\alpha)/2} \le \frac{1}{2}\|v\|_{(1+\alpha)/2}$ whenever $\lambda \ge \lambda_4(\eta, M)$ and $\|f\|_\eta \le M$. As $\|v\|_{(1+\alpha)/2} < \infty$, this implies that $\|v\|_{(1+\alpha)/2} = 0$, and hence $v = 0$. This establishes uniqueness of the solutions to (2.4). $\square$

PROOF OF PROPOSITION 2.7(ii). Let $(f_n)_{n\in\mathbb{Z}_+}, (g_n)_{n\in\mathbb{Z}_+}$ be arbitrary sequences of functions such that $f_n \to f$ and $g_n \to g$ in $\mathcal{C}^{\eta-}$ as $n \to \infty$. Without loss of generality, we can assume that $\sup_{n\in\mathbb{Z}_+}\|f_n\|_\eta \le 2\|f\|_\eta \le 2M$ and $\sup_{n\in\mathbb{Z}_+}\|g_n\|_\eta \le 2\|g\|_\eta \le 2M$. Denote $u := u^\lambda_{f,g}$, $u_n := u^\lambda_{f_n,g_n}$, $v_n := u - u_n$. By part (i), $v_n \in \mathcal{C}^{(1+\alpha)/2}$ and solves

$$\lambda v_n - \mathcal{L}_\alpha v_n - f v_n' = g - g_n + u_n'(f - f_n).$$

Fix arbitrary $\eta' \in (\frac{1-\alpha}{2}, \eta)$. Then we apply Proposition 2.7(i) with η' instead of η to get that for any $\lambda \ge \lambda_0(\eta', 2M)$ we have

$$\begin{aligned}
&\|u_n - u\|_{(1+\alpha)/2} \\
&\quad = \|v_n\|_{(1+\alpha)/2} \\
&\quad \le C\lambda^{-\frac{1}{2}+\frac{1}{2\alpha}-\frac{\eta'}{\alpha}}\big(\|g_n - g\|_{\eta'} + \|f_n - f\|_{\eta'}\|u_n\|_{(1+\alpha)/2}\big)\big(1 + \|f\|_{\eta'}\big) \\
&\quad \le C(1+M)^2\lambda^{-\frac{1}{2}+\frac{1}{2\alpha}-\frac{\eta'}{\alpha}}\big(\|g_n - g\|_{\eta'} + \|f_n - f\|_{\eta'}\|g_n\|_\eta\big) \to 0.
\end{aligned}$$

$\square$

5. Proof of Theorem 2.3. In this section, we present the proof of our main result, Theorem 2.3. We will follow the sketch of the proof presented in Section 2.2. We will rely on the machinery related to the resolvent equation developed in Section 4 and Preliminaries from Section 3. Recall that we have fixed $\alpha \in (1,2)$, $\beta \in (1/2-\alpha/2, 0)$, $b \in \mathcal{C}^\beta$, the initial condition $x \in \mathbb{R}$ and the length of time interval $T > 0$. Our goal is to show that (1.1) has a unique strong solution on time interval $[0, T]$.

We begin with a very standard calculation of the second moment of a stochastic integral. We will use this result a couple of times, and hence for the sake of completeness we decided to state it precisely.

LEMMA 5.1. *Let L be an α-stable Lévy process, ν be its Lévy measure and $\widetilde{N}$ be the compensated Poisson measure associated with L. Assume that $f\colon [0,T] \times \mathbb{R}_+ \times \Omega \to \mathbb{R}$ is a measurable function adapted to the filtration of L. Suppose that there exist $\gamma \in (\alpha/2, 1]$ and constant $C_f > 0$ such that* P-*a.s.*

$$\big|f(s,r,\omega)\big| \le C_f\big(|r|^\gamma \wedge 1\big), \quad s \in [0,T], r \in \mathbb{R}. \tag{5.1}$$

Then there exists a constant $C > 0$ such that for any stopping times $\tau_1, \tau_2 \in [0,T]$ with $\tau_1 \le \tau_2$ we have

$$\mathsf{E}\biggl(\int_{\tau_1}^{\tau_2}\int_{\mathbb{R}} f(s,r,\omega)\widetilde{N}(ds,dr)\biggr)^2 \le CC_f^2\mathsf{E}|\tau_2-\tau_1|. \tag{5.2}$$

The proof of the lemma is given in the Supplementary Material [3], Section 6.

The upcoming subsections are devoted to the proofs of Propositions 2.8, 2.9, 2.10. We complete the proof of Theorem 2.3 in Section 5.4.

5.1. *Proof of Proposition* 2.8: *Any weak solution of* (1.1) *solves the Zvonkin equation*. We will use different properties of integrals with respect to the Dirichlet processes established in Section 3.3. We begin with the following simple moment bound.

LEMMA 5.2. *Let X be a weak solution of SDE* (1.1) *in the class $\mathcal{V}((1+\frac{\beta}{\alpha}) \wedge 1)$. Then for any $\gamma \in [0,\alpha)$ there exists a constant $C > 0$ such that for any $s, t \in [0,T]$,*

$$\mathsf{E}|X_t - X_s|^\gamma \le C|t-s|^{\gamma/\alpha}. \tag{5.3}$$

PROOF. First, note that by basic properties of an α–stable process we have for $\gamma \in [0,\alpha)$,

$$\mathsf{E}|L_t - L_s|^\gamma \le C|t-s|^{\gamma/\alpha}, \quad s,t \in [0,T].$$

It follows from our assumptions that $1/\alpha < 1 + \beta/\alpha$. Therefore, using Jensen's inequality and the fact that X is in the class $\mathcal{V}((1+\frac{\beta}{\alpha}) \wedge 1)$ we obtain for any $s, t \in [0,T]$,

$$\mathsf{E}|X_t - X_s|^\gamma \le C\mathsf{E}|A_t - A_s|^\gamma + C\mathsf{E}|L_t - L_s|^\gamma \le C|t-s|^{\gamma/\alpha}.$$

This immediately yields (5.3). □

LEMMA 5.3 (Krylov-type estimate). *Let X be a weak solution of SDE* (1.1) *in the class $\mathcal{V}((1+\frac{\beta}{\alpha}) \wedge 1)$. Then for any $\delta \in [0, \frac{1}{1/2-\beta/\alpha})$, $\kappa \in [0, 1+\beta/\alpha)$ there exists a constant $C = C(T) > 0$ such that for any $f \in \mathcal{C}_b^\infty$, $0 \le s \le t \le T$ we have*

$$\mathsf{E}\biggl|\int_s^t f(X_l)\,dl\biggr|^\delta \le C|t-s|^{\delta\kappa}\|f\|_\beta^\delta. \tag{5.4}$$

PROOF. We begin by observing that X is a Dirichlet process (as A has zero energy due to (2.2) since X is in the class $\mathcal{V}((1+\frac{\beta}{\alpha})\wedge 1))$. Let us fix $f\in\mathcal{C}_b^\infty$ and $0\le s<t\le T$. Take now any $\delta<1/(1/2-\beta/\alpha)$. Since we took $\beta<0$, we have $\delta<2$.

We claim that it is sufficient to show (5.4) only for those s,t that are close enough; further, we assume that $|t-s|\le 1$. Indeed, if this is already proven, then for any $s,t\in[0,T]$, $s\le t$ we can take an increasing sequence $(t_i)_{i\in[0,N]}$ such that $t_0=s$, $t_N=t$, $t_{i+1}-t_i\le 1$ and $N\le T$. Then

$$\begin{aligned}\mathsf{E}\left|\int_s^t f(X_l)\,dl\right|^\delta &\le N\sum_{i=0}^{N-1}\mathsf{E}\left|\int_{t_i}^{t_{i+1}} f(X_l)\,dl\right|^\delta \le CN\|f\|_\beta^\delta\sum_{i=0}^{N-1}|t_{i+1}-t_i|^{\delta\kappa}\\ &\le C\|f\|_\beta^\delta N^{1+\delta\kappa}|t-s|^{\delta\kappa}\\ &\le C\|f\|_\beta^\delta|t-s|^{\delta\kappa},\end{aligned}$$

for some $C=C(T)$. Thus we can safely assume that $|t-s|\le 1$.

Now let us consider a function $v^\lambda:=u_{0,f}^\lambda$, where $\lambda\ge 1$; this function is well defined by Lemma 4.1. We apply Itô's formula for Dirichlet processes [14], Theorem 3.4 (see also [4], Theorem 5.15(ii)) to derive for $0\le s<t\le T$

$$\begin{aligned}v^\lambda(X_t)-v^\lambda(X_s)&=\int_s^t\mathcal{L}_\alpha v^\lambda(X_l)\,dl+\int_s^t\int_{\mathbb{R}}\big[v^\lambda(X_{l-}+r)-v^\lambda(X_{l-})\big]\widetilde{N}(dl,dr)\\ &\quad+\int_s^t(v^\lambda)'(X_{l-})\,dA_l\\ &=\int_s^t\int_{\mathbb{R}}\big[v^\lambda(X_{l-}+r)-v^\lambda(X_{l-})\big]\widetilde{N}(ds,dr)\\ &\quad+\lambda\int_s^t v^\lambda(X_l)\,dl-\int_s^t f(X_l)\,dl\\ &\quad+\int_s^t(v^\lambda)'(X_l)\,dA_l,\end{aligned}$$

where we also used the fact that v^λ solves (2.4) with 0 in place of f and f in place of g. By rearranging the terms, we get

$$\begin{aligned}\mathsf{E}\left|\int_s^t f(X_l)\,dl\right|^\delta &\le C\|v^\lambda\|^\delta\big(2+\lambda(t-s)\big)^\delta\\ &\quad+C\mathsf{E}\left|\int_s^t\int_{\mathbb{R}}\big[v^\lambda(X_{l-}+r)-v^\lambda(X_{l-})\big]\widetilde{N}(ds,dr)\right|^\delta\\ &\quad+C\mathsf{E}\left|\int_s^t(v^\lambda)'(X_l)\,dA_l\right|^\delta\\ &=I_1+I_2+I_3.\end{aligned}\tag{5.5}$$

Now recall that we supposed that $|t-s|\le 1$. Then we can take $\lambda:=(t-s)^{-1}$. We immediately get by Lemma 4.1 with $\gamma=\varepsilon\alpha/\delta$, $\eta=\beta$ that for any $\varepsilon>0$,

$$I_1\le C(t-s)^{\delta+\delta\beta/\alpha-\varepsilon}\|f\|_\beta^\delta.\tag{5.6}$$

Note that for any $\rho\in(0,1)$

$$\big|v^\lambda(X_{l-}+r)-v^\lambda(X_{l-})\big|\le 2\|v^\lambda\|_\rho\big(|r|^\rho\wedge 1\big).$$

Therefore, we take $\rho = \alpha/2 + \alpha\varepsilon/\delta$ and apply consequently Lemma 5.1, Jensen's inequality, and Lemma 4.1 with $\gamma = \rho$, $\eta = \beta$. We get

$$I_2 \le C(\|v^\lambda\|_\rho)^\delta |t-s|^{\delta/2} \le C|t-s|^{\delta+\delta\beta/\alpha-\varepsilon}\|f\|_\beta^\delta. \tag{5.7}$$

Thus it remains to estimate I_3. Let

$$\varepsilon > 0, \qquad \rho \in (-\beta + \varepsilon\alpha, \alpha + \beta - 1), \qquad \sigma \in [2, \alpha/\rho) \tag{5.8}$$

be parameters to be chosen later. Lemma 4.1 and Lemma 5.2 imply for any $t_1, t_2 \in [0, T]$,

$$\mathsf{E}|(v^\lambda)'(X_{t_1}) - (v^\lambda)'(X_{t_2})|^\sigma \le (\|(v^\lambda)'\|_\rho)^\sigma \mathsf{E}|X_{t_1} - X_{t_2}|^{\rho\sigma} \le C_1\|f\|_\beta^\sigma |t_1 - t_2|^{\rho\sigma/\alpha}$$

for some $C_1 > 0$. Recall that since X is in the class $\mathcal{V}((1 + \frac{\beta}{\alpha}) \wedge 1)$, there exists $C_2 > 0$ such that for any $t_1, t_2 \in [0, T]$,

$$\mathsf{E}|A_{t_1} - A_{t_2}|^2 \le C_2|t_1 - t_2|^{2(1+\beta/\alpha-\varepsilon)}.$$

Thus we can apply Lemma 3.9 with $C_f = C_1^{1/\sigma}\|f\|_\beta$, $C_A = C_2^{1/2}$, $p_1 = \sigma$, $p_2 = 2$, $\gamma_1 = \rho/\alpha$, $\gamma_2 = 1 + \beta/\alpha - \varepsilon$. Our choice of parameters ρ and σ automatically implies that $\gamma_1 + \gamma_2 > 1$ and $1/p_1 + 1/p_2 \le 1$. Thus all the conditions of Lemma 3.9 are satisfied. Further, we also have that

$$\frac{1}{h} := \frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{\sigma} + \frac{1}{2} > 1/2 - \beta/\alpha + \varepsilon.$$

Recall that σ and ρ were arbitrary parameters that satisfy bounds in (5.8). By choosing ρ close enough to its lower bound $-\beta + \varepsilon\alpha$ and choosing σ close enough to its upper bound α/ρ, one can make $1/h$ to be arbitrarily close (though still bigger) to $1/2 - \beta/\alpha + \varepsilon$. Therefore, bound (3.9) yields for any $h < (1/2 - \beta/\alpha + \varepsilon)^{-1}$,

$$\mathsf{E}\left|\int_s^t v'(X_l)\,dA_l\right|^h \le C\|f\|_\beta^h |t-s|^{h(1+\beta/\alpha-\varepsilon)}.$$

Combining this with (5.6) and (5.7) and substituting them into (5.5), we obtain (5.4). This completes the proof of the lemma. □

REMARK 5.4. If one is interested to show existence of the local time for weak solutions to (1.1), one would need to use the Krylov-type estimate presented in (5.4) for a sequence of functions that approximate the delta function. Using the appropriate Hölder–Besov regularity properties of the delta function, one should be able to derive all the details.

PROOF OF PROPOSITION 2.8. Let X be a weak solution to (1.1) in the class $\mathcal{V}((1 + \frac{\beta}{\alpha}) \wedge 1)$. Let $(b_n)_{n\in\mathbb{Z}_+}$ be a sequence of $\mathcal{C}_b^\infty$ functions converging to b in $\mathcal{C}^{\beta-}$. Without loss of generality, we can assume that for each $n \in \mathbb{Z}_+$ we have $\|b_n\|_\beta \le 2\|b\|_\beta$.

Fix $\lambda \ge \lambda_0(\beta, 2\|b\|_\beta)$. Let $u_n = u^\lambda_{b_n,b_n}$ be a unique $\mathcal{C}^{1/2+\alpha/2}$ solution to (2.4). It follows from Proposition 2.7(i) that $u_n \in \mathcal{C}_b^\infty$. Definition 2.1 implies that A has zero energy, and thus X is a Dirichlet process. We apply Itô's formula for Dirichlet processes ([14], Theorem 3.4; see also [4], Theorem 5.15(ii)) to derive for $t \ge 0$,

$$\begin{aligned} u_n(X_t) = u_n(x) &+ \int_0^t \mathcal{L}_\alpha u_n(X_s)\,ds \\ &+ \int_0^t \int_{\mathbb{R}} [u_n(X_{s-} + r) - u_n(X_{s-})]\widetilde{N}(ds, dr) \\ &+ \int_0^t u_n'(X_{s-})\,dA_s. \end{aligned} \tag{5.9}$$

We continue (5.9) as follows, using the fact that u_n solves (2.4):

$$\begin{aligned} u_n(X_t) = u_n(x) &+ \int_0^t \int_{\mathbb{R}} [u_n(X_{s-}+r) - u_n(X_{s-})] \widetilde{N}(ds, dr) \\ &+ \lambda \int_0^t u_n(X_s)\, ds - \int_0^t b_n(X_s)\, ds \\ &+ \int_0^t u_n'(X_s)\, dA_s - \int_0^t u_n'(X_s) b_n(X_s)\, ds. \end{aligned} \tag{5.10}$$

For any fixed $t \in [0, T]$, let us pass to the limit in (5.10) as $n \to \infty$. Since by Proposition 2.7(ii) u_n converges to u in $\mathcal{C}^{1/2+\alpha/2}$, it is clear that

$$\begin{aligned} &u_n(X_t) \to u(X_t), \qquad u_n(x) \to u(x), \\ &\lambda \int_0^t u_n(X_s)\, ds \to \lambda \int_0^t u(X_s)\, ds \quad \text{a.s. as } n \to \infty. \end{aligned} \tag{5.11}$$

Note that since $1/2 + \alpha/2 > 1$, there exists $C > 0$

$$\bigl|u_n(X_{s-}+r) - u(X_{s-}+r) - u_n(X_{s-}) + u(X_{s-})\bigr| \le C \|u_n - u\|_{1/2+\alpha/2} \bigl(|r| \wedge 1\bigr).$$

Therefore by Lemma 5.1 we have

$$\begin{aligned} &\mathsf{E}\biggl(\int_0^t \int_{\mathbb{R}} [u_n(X_{s-}+r) - u(X_{s-}+r) - u_n(X_{s-}) + u(X_{s-})] \widetilde{N}(ds, dr)\biggr)^2 \\ &\quad \le C \|u_n - u\|^2_{1/2+\alpha/2} T. \end{aligned} \tag{5.12}$$

Using again that $\|u_n - u\|_{1/2+\alpha/2} \to 0$ as $n \to \infty$, we deduce from (5.12) that

$$\begin{aligned} &\int_0^t \int_{\mathbb{R}} [u_n(X_{s-}+r) - u_n(X_{s-})] \widetilde{N}(ds, dr) \\ &\quad \to \int_0^t \int_{\mathbb{R}} [u(X_{s-}+r) - u(X_{s-})] \widetilde{N}(ds, dr), \end{aligned} \tag{5.13}$$

in probability as $n \to \infty$.

By the definition of a solution,

$$\int_0^t b_n(X_s)\, ds \to A_t, \tag{5.14}$$

in probability as $n \to \infty$.

Thus, it remains to find the limit of the last two terms in the right-hand side of (5.10). Fix $\varepsilon > 0$ small enough. Applying Proposition 2.7(ii) and Lemma 5.2, we obtain for any $t_1, t_2 \in [0, T]$,

$$\mathsf{E}\bigl|u_n'(X_{t_1}) - u_n'(X_{t_2})\bigr|^\sigma \le \bigl(\|u_n'\|_\rho\bigr)^\sigma \mathsf{E}|X_{t_1} - X_{t_2}|^{\rho\sigma} \le C_1 \|b\|^\sigma_\beta |t_1 - t_2|^{\rho\sigma/\alpha}$$

whenever

$$\rho \in (-\beta + \varepsilon\alpha, \alpha + \beta - 1); \qquad \sigma \in (0, \alpha/\rho). \tag{5.15}$$

Furthermore, by the definition of the solution and Lemma 5.3, for any $t_1, t_2 \in [0, T]$,

$$\begin{aligned} \mathsf{E}|A_{t_1} - A_{t_2}|^h &\le C_2 |t_1 - t_2|^{h(1+\beta/\alpha-\varepsilon)}; \\ \mathsf{E}\biggl|\int_{t_1}^{t_2} b_n(X_l)\, dl\biggr|^h &\le C_2 \|b\|^h_\beta |t_1 - t_2|^{h(1+\beta/\alpha-\varepsilon)}, \end{aligned}$$

whenever $0 \le h < 1/(1/2 - \beta/\alpha)$. Now we can apply Lemma 3.10 to the functions (u_n'), (b_n), A, with the following parameters: $C_{f_1} := C_1^{1/\sigma}\|b\|_\beta$, $p_1 := \sigma$, $\gamma_1 := \rho/\alpha$, $C_A := C_2^{1/h}$, $p_2 = h$, $\gamma_2 = 1 + \beta/\alpha - \varepsilon$. It follows from (5.15) that $\gamma_1 + \gamma_2 > 1$. Note that

$$\frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{\sigma} + \frac{1}{h} > \frac{\rho}{\alpha} + 1/2 - \beta/\alpha > 1/2 - 2\beta/\alpha + \varepsilon.$$

Further, by choosing h close enough to $1/(1/2 - \beta/\alpha)$, ρ close enough to $-\beta + \varepsilon\alpha$, σ close enough to α/ρ, one can guarantee that $\frac{1}{p_1} + \frac{1}{p_2}$ will be arbitrarily close to $1/2 - 2\beta/\alpha + \varepsilon$. However, for ε small enough we have $1/2 - 2\beta/\alpha + \varepsilon < 1$. Hence for some suitable choice of parameters, one has $\frac{1}{p_1} + \frac{1}{p_2} < 1$. Hence all the conditions of Lemma 3.10 are satisfied. Thus

$$\int_0^t u_n'(X_s)\, dA_s - \int_0^t u_n'(X_s) b_n(X_s)\, ds \to 0,$$

in probability as $n \to \infty$. Combining this with (5.11), (5.13), (5.14), we can pass to the limit in probability in (5.10) as $n \to \infty$. We obtain that for each fixed t the following identity holds a.s.:

$$\begin{aligned} u(X_t) = u(x) &+ \int_0^t \int_{\mathbb{R}} [u(X_{s-} + r) - u(X_{s-})] \widetilde{N}(ds, dr) \\ &+ \lambda \int_0^t u(X_s)\, ds - A_t. \end{aligned}$$

To complete the proof, it remains to note that $A_t = X_t - L_t - x$; thus X_t is indeed a weak solution to equation (2.7). $\square$

5.2. *Proof of Proposition* 2.9: *Weak existence*. In this section, we establish Proposition 2.9.

As explained in Section 2.1, we will construct a sequence of solutions to the approximated equations with smooth coefficients and then prove that this sequence has a limiting point, which solves SDE (1.1) in the weak sense. Thus let $(b_n)_{n\in\mathbb{Z}_+}$ be a sequence of $\mathcal{C}_b^\infty$ functions converging to b in $\mathcal{C}^{\beta-}$. Without loss of generality, we can assume that for each $n \in \mathbb{Z}_+$ we have $\|b_n\|_\beta \le 2\|b\|_\beta$. Recall the definitions of $(X^n)_{n\in\mathbb{Z}_+}$ and $(A^n)_{n\in\mathbb{Z}_+}$, which are given in (2.8) and (2.9), correspondingly. Recall the definition of the function λ_0 in Proposition 2.7.

For $\lambda \ge \lambda_0(\beta, 2\|b\|_\beta)$, let $u_n^\lambda := u_{b_n,b_n}^\lambda$ be the unique solution of the resolvent equation (2.4) in class $\mathcal{C}^{\frac{1+\alpha}{2}}$. By Proposition 2.7(i), u_n^λ is well defined and $u_n^\lambda \in \mathcal{C}_b^\infty$. For brevity, in this subsection further we will write just λ_0 instead of $\lambda_0(\beta, 2\|b\|_\beta)$.

LEMMA 5.5.

(i) *For each* $\lambda \ge \lambda_0$, $n \in \mathbb{Z}_+$, $t \in [0, T]$, *we have*

$$\begin{aligned} u_n^\lambda(X_t^n) = u_n^\lambda(x) &+ \int_0^t \int_{\mathbb{R}} (u_n^\lambda(X_{s-}^n + r) - u_n^\lambda(X_{s-}^n)) \widetilde{N}(ds, dr) \\ &+ \lambda \int_0^t u_n^\lambda(X_s^n)\, ds - A_t^n. \end{aligned} \tag{5.16}$$

(ii) *Further, for any* $\varepsilon > 0$ *there exists a constant* $C > 0$ *such that for any* $n \in \mathbb{Z}_+$, $\delta \le 1/\lambda_0$, *and stopping time* $\tau \in [0, T]$ *we have*

$$\mathsf{E}|A_{\tau+\delta}^n - A_\tau^n|^2 \le C\delta^{2(1+\frac{\beta}{\alpha}-\varepsilon)}\|b\|_\beta^2(\|b\|_\beta + 1)^2. \tag{5.17}$$

(iii) *There exists a constant* $C > 0$ *such that for any* $n \in \mathbb{Z}_+$, $t \in [0, T]$ *we have*

$$\mathsf{E}|A_t^n|^2 \le C\|b\|_\beta^2(\|b\|_\beta + 1)^2. \tag{5.18}$$

(iv) *Finally, the sequence* $\{(X^n, A^n)\}_{n\in\mathbb{Z}_+}$ *is tight in* $\mathbb{D}_{\mathbb{R}^2}[0, T]$.

PROOF. (i) Since $u_n^\lambda \in \mathcal{C}_b^\infty$, identity (5.16) follows immediately by an application of Itô's formula (see, e.g., [2], Theorem 4.4.7) to the process X^n and the function u_n^λ.

(ii) Fix $\delta \le 1/\lambda_0$, stopping time τ, and $\varepsilon > 0$ small enough. For $\lambda \ge \lambda_0$, $n \in \mathbb{Z}_+$ we denote

$$I_t^{n,\lambda} := \int_0^t \int_{\mathbb{R}} (u_n^\lambda(X_{s-}^n + r) - u_n^\lambda(X_{s-}^n))\widetilde{N}(ds, dr), \quad t \in [0, T].$$

It follows from (5.16) that for any $\lambda \ge \lambda_0$ we have

$$\begin{aligned} |A_{\tau+\delta}^n - A_\tau^n| &\le |u_n^\lambda(X_{\tau+\delta}^n) - u_n^\lambda(X_\tau^n)| + |I_{\tau+\delta}^{n,\lambda} - I_\tau^{n,\lambda}| \\ &\quad + \lambda \int_\tau^{\tau+\delta} |u_n^\lambda(X_s^n)|\,ds \\ &\le \|u_n^\lambda\|(2 + \lambda\delta) + |I_{\tau+\delta}^{n,\lambda} - I_\tau^{n,\lambda}|. \end{aligned} \tag{5.19}$$

Now let us pick $\lambda := \delta^{-1}$. Since $\delta \le 1/\lambda_0$, we clearly have $\lambda \ge \lambda_0$. Then it follows from Proposition 2.7(i) with $\gamma = \varepsilon\alpha$, $\eta = \beta$ and the bound $\|b_n\|_\beta \le 2\|b\|_\beta$ that

$$\|u_n^\lambda\|(2 + \lambda\delta) \le C\delta^{1+\frac{\beta}{\alpha}-\varepsilon}\|b\|_\beta(\|b\|_\beta + 1), \tag{5.20}$$

where the constant $C > 0$ depends only on α, β and ε (but not n, δ, λ or τ).

Note that for any $\gamma \in (0, 1)$, we have

$$|u_n^\lambda(X_{s-}^n + r) - u_n^\lambda(X_{s-}^n)| \le \|u_n^\lambda\|_\gamma(1 \wedge |r|^\gamma), \quad s \in [0, T], r \in \mathbb{R}.$$

Thus we can take $\gamma := \alpha/2 + \alpha\varepsilon$ and deduce from Lemma 5.1 and Proposition 2.7(i) with $\gamma = \alpha/2 + \varepsilon\alpha$, $\eta = \beta$ that

$$\mathsf{E}(I_{\tau+\delta}^{n,\lambda} - I_\tau^{n,\lambda})^2 \le C\delta(\|u_n^\lambda\|_{\alpha/2+\alpha\varepsilon})^2 \le C\delta^{2(1+\frac{\beta}{\alpha}-\varepsilon)}\|b\|_\beta^2(\|b\|_\beta + 1)^2,$$

where again the constant C does not depend on n, δ, λ or τ. Combining this bound with (5.19) and (5.20), we establish (5.17).

(iii) It is clear that for any $t \in [0, T]$ there exists $N \in \mathbb{Z}_+$ and a increasing sequence $(t_i)_{i\in[0,N]}$ such that $t_0 = 0$, $t_N = t$ and $t_{i+1} - t_i \le 1/\lambda_0$. Further, one can take $N = \lceil T\lambda_0 \rceil$. Then it follows from part (ii) of the lemma and the fact that $A_0 = 0$ that

$$\mathsf{E}|A_t^n|^2 \le N \sum_{i=0}^{N-1} \mathsf{E}|A_{t_{i+1}}^n - A_{t_i}^n|^2 \le C(T, \lambda_0, \alpha, \beta)\|b\|_\beta^2(\|b\|_\beta + 1)^2,$$

which proves (5.18).

(iv) To establish the tightness of $\{(X^n, A^n)\}_{n\in\mathbb{Z}_+}$, first let us verify that the sequence $(A^n)_{n\in\mathbb{Z}_+}$ is tight in $\mathbb{D}_{\mathbb{R}}[0, T]$. We would like to apply the Aldous theorem [1], Theorem 1. Thus we need to check that for each $t \in [0, T]$ the sequence of random variables $(A_t^n)_{n\in\mathbb{Z}_+}$ is tight; and that for any sequence of stopping times $(\tau_n)_{n\in\mathbb{Z}_+}$ and constants $\delta_n \to 0$ we have

$$A_{\tau_n+\delta_n}^n - A_{\tau_n}^n \to 0 \quad \text{as } n \to \infty \text{ in probability.} \tag{5.21}$$

The first condition of Aldous' theorem holds thanks to part (iii) of the lemma. Indeed, bound (5.18) yields that for each fixed $t \in [0, T]$ the sequence $(A_t^n)_{n\in\mathbb{Z}_+}$ is tight.

To verify the second condition of Aldous' theorem, we take a sequence of stopping times $(\tau_n)_{n\in\mathbb{Z}_+}$ and a sequence of constants $\delta_n \to 0$. We can assume without loss of generality that $\delta_n \le 1/\lambda_0$ for all $n \in \mathbb{Z}_+$. Then we apply part (ii) of the lemma with $\tau = \tau_n$, $\delta = \delta_n$. We derive

$$\mathsf{E}\big|A^n_{\tau_n+\delta_n} - A^n_{\tau_n}\big| \le C\|b\|_\beta\big(\|b\|_\beta + 1\big)\sqrt{\delta_n} \to 0 \quad \text{as } n \to \infty, \tag{5.22}$$

where we used the fact that $1 + \beta/\alpha > 1/2$. This implies (5.21). Thus all the conditions of [1], Theorem 1, are satisfied and the sequence $(A^n)_{n\in\mathbb{Z}_+}$ is tight.

Recall that $X^n_t = A^n_t + x + L_t$. It follows from (5.18) and (5.22) that the sequence $(X^n)_{n\in\mathbb{Z}_+}$ also satisfies the conditions of [1], Theorem 1. Hence $(X^n)_{n\in\mathbb{Z}_+}$ is tight.

To complete the proof, it remains to note that A^n is continuous in t for each n. Thus $(A^n)_{n\in\mathbb{Z}_+}$ is actually $\mathcal{C}$-tight and, therefore, the sequence $\{(X^n, A^n)\}_{n\in\mathbb{Z}_+}$ is tight in $\mathbb{D}_{\mathbb{R}^2}[0, T]$ by [27], Corollary VI.3.33(b). $\square$

Now we are ready to prove the main result of this subsection.

PROOF OF PROPOSITION 2.9(i). Fix $\lambda \ge \lambda_0(\beta, 2\|b\|_\beta)$. In this proof, for brevity we will write $u := u^\lambda_{b,b}$ and $u_n := u^\lambda_{b_n,b_n}$.

We use the approximating sequence $\{(X^n, A^n)\}_{n\in\mathbb{Z}_+}$ constructed in Lemma 5.5. It follows from Lemma 5.5 that this sequence is tight in $\mathbb{D}_{\mathbb{R}^2}[0, T]$. Hence by the Prokhorov theorem there exists a subsequence (n_k) such that (X^{n_k}, A^{n_k}) converges weakly in the Skorokhod space $\mathbb{D}_{R^2}[0, T]$ to the limit (X, A). In order not to overburden the notation, we suppose that we have already passed to this subsequence, and thus we assume that (X^n, A^n) converges weakly to (X, A). Then by the Skorokhod representation theorem (see, e.g., [16], Theorem 3.1.8) there exists a sequence of random elements $(\widehat{X}^n, \widehat{A}^n)$ defined on a common probability space $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathsf{P}})$ such that $(\widehat{X}^n, \widehat{A}^n) \to (\widehat{X}, \widehat{A})$ a.s. in the Skorokhod metric and $\operatorname{Law}(\widehat{X}^n, \widehat{A}^n) = \operatorname{Law}(X^n, A^n)$, $\operatorname{Law}(\widehat{X}, \widehat{A}) = \operatorname{Law}(X, A)$.

Denote $\widehat{L}^n := \widehat{X}^n - \widehat{A}^n - x$, and define similarly $\widehat{L}$. By the above,

$$\operatorname{Law}(\widehat{L}^n) = \operatorname{Law}(X^n - A^n - x) = \operatorname{Law}(L).$$

Thus, $\widehat{L}^n$ is an α-stable Lévy process. It follows from Lemma 3.3, that $(\widehat{X}^n, \widehat{A}^n, \widehat{L}^n)$ converges a.s. in the Skorokhod metric to $(\widehat{X}, \widehat{A}, \widehat{L})$. Hence $\widehat{L}$ is also an α-stable Lévy process. Denote by $\widetilde{\widehat{N}}{}^n$ (resp., $\widetilde{\widehat{N}}$) the compensated Poisson random measure of $\widehat{L}^n$ (resp., $\widehat{L}$).

It follows from the above considerations and (5.16) that for $t \in [0, T]$,

$$\begin{aligned} &u_n(\widehat{X}^n_t) - u_n(x) - \lambda\int_0^t u_n(\widehat{X}^n_s)\,ds + \widehat{A}^n_t \\ &\quad = \int_0^t\int_{\mathbb{R}} \big(u_n(\widehat{X}^n_{s-} + r) - u_n(\widehat{X}^n_{s-})\big)\widetilde{\widehat{N}}{}^n(ds, dr). \end{aligned} \tag{5.23}$$

Let us pass to the limit as $n \to \infty$ in (5.23).

First, we recall that $\widehat{X}^n$ converges a.s. to $\widehat{X}$ in the Skorokhod metric as $n \to \infty$. By Proposition 2.7(i), we have $\sup_n \|u_n\|_{(1+\alpha)/2} < \infty$ and by Proposition 2.7(ii) $\lim_{n\to\infty} \|u_n - u\|_{(1+\alpha)/2} \to 0$. Therefore, by Lemma 3.3,

$$u_n(\widehat{X}^n_\cdot) \to u(\widehat{X}_\cdot) \quad \text{as } n \to \infty \text{ a.s. in } \mathbb{D}_{\mathbb{R}}[0, T]. \tag{5.24}$$

By Lemma 3.4,

$$\lambda\int_0^\cdot u_n(\widehat{X}^n_s)\,ds \to \lambda\int_0^\cdot u(\widehat{X}_s)\,ds \quad \text{as } n \to \infty \text{ a.s. in } \mathbb{D}_{\mathbb{R}}[0, T]. \tag{5.25}$$

Note that the function $\int_0^\cdot u(\widehat{X}_s)\,ds$ is continuous; recall that $\widehat{A}^n$ converges a.s. to a continuous function $\widehat{A}$. Therefore, (5.24), (5.25) and [27], Proposition VI.1.23, yield that the left-hand side of (5.23) converges a.s. in $\mathbb{D}[0,T]$ to

$$u(\widehat{X}_\cdot) - u(x) - \lambda \int_0^\cdot u(\widehat{X}_s)\,ds + \widehat{A}.$$

Recall that $(\widehat{X}^n, \widehat{L}^n)$ converges a.s. in $\mathbb{D}_{\mathbb{R}^2}[0,T]$ to $(\widehat{X}, \widehat{L})$ and $\sup_n \|u_n\|_{(1+\gamma)/2} < \infty$. As $\widehat{X}^n$ is a strong solution, it is adapted to the filtration generated by $\widetilde{\widehat{N}^n}$, and thus all the conditions of Lemma 3.5 are satisfied. Hence the right-hand side of (5.23) converges in probability in $\mathbb{D}[0,T]$ to

$$\int_0^\cdot \int_{\mathbb{R}} \big(u(\widehat{X}_{s-}+r) - u(\widehat{X}_{s-})\big)\widetilde{\widehat{N}}(ds,dr).$$

Thus,

$$\begin{aligned}
&u(\widehat{X}_t) - u(x) - \lambda \int_0^t u(\widehat{X}_s)\,ds + \widehat{A}_t \\
&\quad = \int_0^t \int_{\mathbb{R}} \big(u(\widehat{X}_{s-}+r) - u(\widehat{X}_{s-})\big)\widetilde{\widehat{N}}(ds,dr), \qquad t \in [0,T].
\end{aligned} \tag{5.26}$$

Since $\widehat{A}_t = \widehat{X}_t - \widehat{L}_t - x$, we see that $(\widehat{X}, \widehat{L})$ is indeed a weak solution to (2.7). $\square$

PROOF OF PROPOSITION 2.9(ii). Recall that by definition

$$\widehat{X}_t = x + \widehat{A}_t + \widehat{L}_t, \qquad t \in [0,T].$$

Thus it remains to check that the process $\widehat{A}_t$ satisfies the second property in Definition 2.1 and $\widehat{X}$ is in the class $\mathcal{V}((1+\frac{\beta}{\alpha}) \wedge 1)$.

To check the second property, take any approximating sequence $(b_n)_{n\in\mathbb{Z}_+} \in \mathcal{C}_b^\infty$ such that $b_n \to b$ in $\mathcal{C}^{\beta-}$ as $n \to \infty$ and $\|b_n\|_\beta \le 2\|b\|_\beta$, $n \in \mathbb{Z}_+$. Take any $\lambda \ge \lambda_0(\beta, 2\|b\|_\beta)$. For any $n, m \in \mathbb{Z}_+$, we consider $u^{n,m} := u^\lambda_{b_n,b_m}$, which is the unique $\mathcal{C}^{(1+\alpha/2)}$ solution to the equation (2.4) with b_n in place of f and b_m in place of g. We apply Itô's formula to the process $\widehat{X}^n$. We get

$$\begin{aligned}
&u^{n,m}(\widehat{X}^n_t) - u^{n,m}(x) - \lambda \int_0^t u^{n,m}(\widehat{X}^n_s)\,ds + \int_0^t b_m(\widehat{X}^n_s)\,ds \\
&\quad = \int_0^t \int_{\mathbb{R}} \big(u^{n,m}(\widehat{X}^n_{s-}+r) - u^{n,m}(\widehat{X}^n_{s-})\big)\widetilde{\widehat{N}^n}(ds,dr).
\end{aligned} \tag{5.27}$$

Consider now $u^{(m)} := u^\lambda_{b,b_m}$. Then by Proposition 2.7(ii),

$$\lim_{n\to\infty} \|u^{n,m} - u^{(m)}\|_{(1+\alpha)/2} = 0.$$

Now for each fixed $m \in \mathbb{Z}_+$ we pass to the limit as $n \to \infty$ in (5.27). Arguing exactly as in in part (i) of the proof, we apply Lemmas 3.3, 3.4, 3.5 and [27], Proposition VI.1.23 to obtain

$$\begin{aligned}
&u^{(m)}(\widehat{X}_t) - u^{(m)}(x) - \lambda \int_0^t u^{(m)}(\widehat{X}_s)\,ds + \int_0^t b_m(\widehat{X}_s)\,ds \\
&\quad = \int_0^t \int_{\mathbb{R}} \big(u^{(m)}(\widehat{X}_{s-}+r) - u^{(m)}(\widehat{X}_{s-})\big)\widetilde{\widehat{N}}(ds,dr).
\end{aligned}$$

Comparing this identity with (5.26), we deduce

$$
\begin{aligned}
&\Big\|\int_0^\cdot b_m(\widehat{X}_s)\,ds-\widehat{A}\Big\| \\
&\quad\le \|u^{(m)}-u\|\Big(2+(C+\lambda)T+2\sum_{s\le T}\mathbb{1}\big(|\Delta L_s|>1\big)\Big)+\|J^m\|,
\end{aligned}
\tag{5.28}
$$

where we denoted

$$
\begin{aligned}
J^m(t):=\int_0^t\int_{|r|<1}&\big(u^{(m)}(\widehat{X}_{s-}+r)-u^{(m)}(\widehat{X}_{s-})\\
&-u(\widehat{X}_{s-}+r)-u(\widehat{X}_{s-})\big)\widetilde{\widetilde{N}}(ds,dr),\quad t\in[0,T].
\end{aligned}
$$

Clearly, for any $x\in\mathbb{R}$, $y\ge 0$,

$$
\begin{aligned}
&\big|u^{(m)}(x+y)-u^{(m)}(x)-u(x+y)-u(x)\big|\\
&\quad\le\int_x^{x+y}\big|(u^{(m)})'(s)-u'(s)\big|\,ds\le y\|u^{(m)}-u\|_{(1+\alpha)/2}.
\end{aligned}
$$

Taking into account this inequality and the fact that J^m is a martingale, we apply Doob's inequality to derive for any $\varepsilon>0$,

$$
\mathsf{P}\big(\|J^m\|>\varepsilon\big)\le\varepsilon^{-2}\mathsf{E}J^m(T)^2\le CT\varepsilon^{-2}\|u^{(m)}-u\|^2_{(1+\alpha)/2}.
\tag{5.29}
$$

By Proposition 2.7(ii), $\|u^{(m)}-u\|_{(1+\alpha)/2}\to 0$ as $m\to\infty$. Hence, combining (5.28) and (5.29), we get

$$
\Big\|\int_0^\cdot b_m(\widehat{X}_s)\,ds-\widehat{A}\Big\|\to 0\quad\text{in probability as } m\to\infty.
$$

It remains to show that $\widehat{X}$ is in the class $\mathcal{V}((1+\frac{\beta}{\alpha})\wedge 1)$. Fix any $0\le s\le t\le T$. By the standard argument, we see that it is enough to check (2.2) only for s, t close enough. Thus we can assume that $|t-s|\le\frac{1}{\lambda_0(\beta,2\|b\|_\beta)}$. It follows from Lemma 5.5(ii) that

$$
\mathsf{E}\big|\widehat{A}^n_t-\widehat{A}^n_s\big|^2=\mathsf{E}\big|A^n_t-A^n_s\big|^2\le C|t-s|^{2(1+\frac{\beta}{\alpha}-\varepsilon)}\|b\|^2_\beta\big(\|b\|_\beta+1\big)^2.
$$

By Fatou's lemma,

$$
\mathsf{E}|\widehat{A}_t-\widehat{A}_s|^2\le C|t-s|^{2(1+\frac{\beta}{\alpha}-\varepsilon)}\|b\|^2_\beta\big(\|b\|_\beta+1\big)^2.
$$

This concludes the proof. $\square$

5.3. *Proof of Proposition* 2.10: *Pathwise uniqueness*. First of all, we note that thanks to Proposition 2.7(i), there exists $\lambda_1=\lambda_1(\beta,\|b\|_\beta)\ge\lambda_0(\beta,\|b\|_\beta)$ such that for any $\lambda>\lambda_1$ we have

$$
\|(u^\lambda_{b,b})'\|+\|u^\lambda_{b,b}\|\le 1/4.
\tag{5.30}
$$

For the rest of this section, we fix $\lambda>\lambda_1$. For brevity, we will write u for $u^\lambda_{b,b}$. Introduce the following functions:

$$
\begin{aligned}
&\phi(w):=w+u(w),\qquad \widetilde{b}(w):=u\big(\phi^{-1}(w)\big),\\
&\widetilde{\sigma}(w,r):=u\big(\phi^{-1}(w)+r\big)-u\big(\phi^{-1}(w)\big),\quad w,r\in\mathbb{R}.
\end{aligned}
$$

We will need the following auxiliary lemma.

LEMMA 5.6. *Let $\gamma_1, \gamma_2 \in [0,1)$. Denote $\gamma := \gamma_1 + \gamma_2$ and assume that $\gamma \neq 1$. Then the following hold*:

(i) *For any $f \in \mathcal{C}^\gamma$, and $r, x_1, x_2 \in \mathbb{R}$ we have*

$$\big|f(x_1+r) - f(x_1) - f(x_2+r) + f(x_2)\big| \le 2\|f\|_\gamma |x_1 - x_2|^{\gamma_1} |r|^{\gamma_2}. \tag{5.31}$$

(ii) *For all $x_1, x_2 \in \mathbb{R}$*,

$$\frac{3}{4}|x_1 - x_2| \le \big|\phi(x_1) - \phi(x_2)\big| \le \frac{5}{4}|x_1 - x_2|.$$

In particular, this implies that $\phi(\cdot)$ is invertible and the inverse is Lipschitz.

(iii) *For all $w_1, w_2, r \in \mathbb{R}$*,

$$\big|\widetilde{b}(w_1) - \widetilde{b}(w_2)\big| \le |w_1 - w_2|, \tag{5.32}$$

$$\big|\widetilde{\sigma}(w_1, r) - \widetilde{\sigma}(w_2, r)\big| \le \left(\frac{2}{3}|w_1 - w_2|\right) \wedge |r| \wedge 1. \tag{5.33}$$

The proof of the above lemma is standard and we prove it in the Supplementary Material [3], Section 7.

Let X_t be a solution to (2.7) and put $W_t := \phi(X_t)$. Then rewriting (2.7), we get

$$W_t = \phi(x) + \lambda \int_0^t \widetilde{b}(W_s)\,ds + \int_0^t \int_{\mathbb{R}} \widetilde{\sigma}(W_{s-}, r)\widetilde{N}(ds, dr) + L_t. \tag{5.34}$$

PROPOSITION 5.7. *Equation (5.34) has a pathwise unique solution.*

PROOF. Let W_t be a solution to (5.34). Using (2.6), we can rewrite (5.34) as

$$\begin{aligned} W_t = \phi(x) &+ \int_0^t \widetilde{b}_0(W_s)\,ds + \int_0^t \int_{\mathbb{R}} \big(\widetilde{\sigma}(W_{s-}, r) + r\big)\mathbb{1}_{|r|<1}\widetilde{N}(ds, dr) \\ &+ \int_0^t \int_{\mathbb{R}} \big(\widetilde{\sigma}(W_{s-}, r) + r\big)\mathbb{1}_{|r|\ge 1} N(ds, dr), \end{aligned} \tag{5.35}$$

where

$$\widetilde{b}_0(w) := \lambda \widetilde{b}(w) - c_\alpha \int_{\mathbb{R}} \widetilde{\sigma}(w, r)\mathbb{1}_{|r|\ge 1}|r|^{-\alpha-1}\,dr.$$

We are now going to apply [36], Theorem 3.2, which will guarantee pathwise uniqueness for equation (5.35). Let us verify that all the hypotheses of this theorem are satisfied.

We begin by observing that thanks to (5.32) the function $\widetilde{b}$ is Lipschitz. We also see that by (5.33), there exists a constant $C > 0$ such that for any $w_1, w_2 \in \mathbb{R}$,

$$\int_{|r|\ge 1} \big|\widetilde{\sigma}(w_1, r) - \widetilde{\sigma}(w_2, r)\big| |r|^{-1-\alpha}\,dr \le C|w_1 - w_2|.$$

Thus condition (3.a) of [36], Theorem 3.2 holds.

It follows from (5.33) that the function $\widetilde{\sigma}$ is Lipschitz in w with the Lipschitz constant less than 1. Therefore, for each $r \in \mathbb{R}$,

$$w \to w + \big(\widetilde{\sigma}(w, r) + r\big) \tag{5.36}$$

is a nondecreasing function of w. Using (5.31) with $\gamma_1 = 1/2$, $\gamma_2 = \alpha/2 + \varepsilon/2$ (and $\epsilon > 0$ sufficiently small such that $\beta + \alpha > \frac{1+\alpha+\varepsilon}{2}$), we deduce for any $r, w_1, w_2 \in \mathbb{R}$,

$$\big|\widetilde{\sigma}(w_1, r) - \widetilde{\sigma}(w_2, r)\big| \le C\|u\|_{\frac{1+\alpha+\varepsilon}{2}} |w_1 - w_2|^{\frac{1}{2}} |r|^{\frac{\alpha+\varepsilon}{2}}.$$

Therefore,

$$\begin{aligned}
&\int_{|r|\le 1} \left|\widetilde{\sigma}(w_1,r)-\widetilde{\sigma}(w_2,r)\right|^2 |r|^{-1-\alpha}\,dr \\
&\quad\le C\|u\|^2_{\frac{1+\alpha+\varepsilon}{2}}|w_1-w_2|\int_{|r|\le 1}|r|^{\alpha+\varepsilon}|r|^{-1-\alpha}\,dr \\
&\quad\le C(\epsilon)\|u\|^2_{\frac{1+\alpha+\varepsilon}{2}}|w_1-w_2|.
\end{aligned} \tag{5.37}$$

Since $\beta > \frac{1-\alpha}{2}$, by Proposition 2.7(i) (take $\eta=\beta$ there and also recall that $\beta+\alpha > \frac{1+\alpha+\varepsilon}{2} > \beta$), we have $\|u\|_{\frac{1+\alpha+\varepsilon}{2}} < \infty$. This, (5.36) and (5.37) implies that condition (3.b) of [36], Theorem 3.2 holds.

Thus all the conditions of [36], Theorem 3.2 are satisfied, and hence pathwise uniqueness for (5.35) holds. Consequently, pathwise uniqueness for (5.34) holds. □

PROOF OF PROPOSITION 2.10. Let X be a solution to (2.7). Then $W_t=\phi(X_t)$ is a solution to (5.34). By Proposition 5.7, pathwise uniqueness holds for (5.34) and by Lemma 5.6(b) the function ϕ is invertible. This implies that pathwise uniqueness holds for (2.7). □

As mentioned in the Introduction, a direct proof of Proposition 2.10 is given in the Supplementary Material, [3], Section 8.

5.4. *Proof of Theorem* 2.3. We have already shown that by Proposition 2.9(ii), (1.1) has a weak solution in the class $\mathcal{V}((1+\frac{\beta}{\alpha})\wedge 1)$. By Proposition 2.10, we know that pathwise uniqueness holds for (2.7) and via Proposition 2.8 we know that every weak solution to (1.1) in the class $\mathcal{V}((1+\frac{\beta}{\alpha})\wedge 1)$ is a weak solution to (2.7). Therefore we have shown pathwise uniqueness for (1.1). Now we can apply a generalized version of the classical Yamada–Watanabe theorem; see [31], Theorem 3.4 and Proposition 2.13. The weak existence and pathwise uniqueness for (1.1) then imply strong existence for the same equation. For the sake of completeness, we rewrite our equation (1.1) in the notation of [31] and verify that all the assumptions required for Theorem 3.4 and Proposition 2.13 in [31] are satisfied.

Let $x\in\mathbb{R}$. Let $(\Omega,\mathcal{F},\mathsf{P})$ be a complete probability space and $(L_t)_{t\ge 0}$ be the symmetric stable α-process on it. Let Υ be the product measure of δ_x and the law of L. Let $S_1 = \mathbb{D}_{\mathbb{R}}[0,\infty)$ and $S_2=\mathbb{R}\times\mathbb{D}_{\mathbb{R}}[0,\infty)$. Let $\mathcal{P}(S_1\times S_2)$ be the space of probability measures on $S_1\times S_2$ with the product Borel σ-algebras of S_1 and S_2. Let X be a random variable on $(\Omega,\mathcal{F},\mathsf{P})$ taking values in S_1, Y be a random variable on $(\Omega,\mathcal{F},\mathsf{P})$ taking values in S_2 and $\mu_{X\times Y}\in\mathcal{P}(S_1\times S_2)$ be the joint distribution of (X,Y). Moreover we assume that $Y=(x,(L_s)_{s\ge 0})$, that is, the law of Y is Υ. Then our model (1.1) is specified by a set of constraints Γ relating (X,Y) where the constraint Γ is given by

$$\Gamma := \Bigl\{ Y=\bigl(x,(L_s)_{s\ge 0}\bigr) \text{ and } X \text{ is a solution to (1.1)} \\ \text{in the class } \mathcal{V}\biggl(\biggl(1+\frac{\beta}{\alpha}\biggr)\wedge 1\biggr)\Bigr\}.$$

We denote by

$$S_{\Gamma,\Upsilon} := \bigl\{\mu_{(X,Y)}\in\mathcal{P}(S_1\times S_2) : X,Y \text{ satisfy } \Gamma \text{ and } \mu_{X,Y}(S_1\times\cdot)=\Upsilon(\cdot)\bigr\}.$$

We will follow [31] in defining the notion of compatible solution.

DEFINITION 5.8 (Compatible solutions). For each $t \geq 0$, let $\{\mathcal{F}_t^X\}$ and $\{\mathcal{F}_t^Y\}$ be complete filtrations generated by X and Y, respectively (see Remark 2.3 in [31] for the precise definition of completion). The collection

$$\mathcal{C} \equiv \{(\mathcal{F}_t^X, \mathcal{F}_t^Y) : t \geq 0\}$$

will be referred to as a compatibility structure. X is said to be $\mathcal{C}$-compatible with Y if for each $t \geq 0$, and $h \in \mathbb{L}^1(S_2, \Upsilon)$,

$$\mathsf{E}\big(h(Y) \mid \mathcal{F}_t^X \vee \mathcal{F}_t^Y\big) = \mathsf{E}\big(h(Y) \mid \mathcal{F}_t^Y\big).$$

Finally, let

$$S_{\Gamma,\mathcal{C},\Upsilon} := \{\mu_{(X,Y)} \in S_{\Gamma,\Upsilon} : X \text{ is } \mathcal{C}\text{-compatible with } Y\}.$$

LEMMA 5.9. *(X, L) is a weak solution of (1.1) in the class $\mathcal{V}((1 + \frac{\beta}{\alpha}) \wedge 1)$ if and only if $\mu_{(X,Y)} \in S_{\Gamma,\mathcal{C},\Upsilon}$.*

PROOF. Let (X, L) be a weak solution to (1.1) in the class $\mathcal{V}((1 + \frac{\beta}{\alpha}) \wedge 1)$ and adapted to a complete filtration $\mathcal{F}_t$. Define $Y_{\cdot\leq t} := (x, (L_{\min\{s,t\}})_{s\geq 0})$ and $Y^{\cdot\geq t} := (L_{t+s} - L_t)_{s\geq 0}$.

Clearly, for each $t \geq 0$ we have $\mathcal{F}_t^Y \subset \mathcal{F}_t^X \vee \mathcal{F}_t^Y \subset \mathcal{F}_t$. Note that $Y_{\cdot\leq t}$ is $\mathcal{F}_t^Y$-measurable. Further, $(L_t)_{t\geq 0}$ is an α-stable process with respect to the filtration $\mathcal{F}_t$, and thus also with respect to $\mathcal{F}_t^X \vee \mathcal{F}_t^Y$. This implies that $Y^{\cdot\geq t}$ is independent of $\mathcal{F}_t^X \vee \mathcal{F}_t^Y$. For any $h \in \mathbb{L}^1(S_2, \Upsilon)$ and for all $t \geq 0$, there exist bounded measurable functions h_t on $\mathbb{R} \times \mathbb{D}_{\mathbb{R}}[0, \infty) \times \mathbb{D}_{\mathbb{R}}[0, \infty)$ such $h(Y) = h_t(Y_{\cdot\leq t}, Y^{\cdot\geq t})$ a.s. Then, following the argument in the proof of Lemma 2.4 in [31], we get

$$\begin{aligned}
\mathsf{E}\big(h(Y) \mid \mathcal{F}_t^X \vee \mathcal{F}_t^Y\big) &= \mathsf{E}\big(h_t\big(Y_{\cdot\leq t}, Y^{\cdot\geq t}\big) \mid \mathcal{F}_t^X \vee \mathcal{F}_t^Y\big) \\
&= \mathsf{E}\bigg(\int h_t(Y_{\cdot\leq t}, y)\mathsf{P}\big(Y^{\cdot\geq t} \in dy\big) \,\bigg|\, \mathcal{F}_t^X \vee \mathcal{F}_t^Y\bigg) \\
&= \int h_t(Y_{\cdot\leq t}, y)\mathsf{P}\big(Y^{\cdot\geq t} \in dy\big) \\
&= \mathsf{E}\bigg(\int h_t(Y_{\cdot\leq t}, y)\mathsf{P}\big(Y^{\cdot\geq t} \in dy\big) \,\bigg|\, \mathcal{F}_t^Y\bigg) \\
&= \mathsf{E}\big(h(Y) \mid \mathcal{F}_t^Y\big).
\end{aligned}$$

Thus X is $\mathcal{C}$-compatible with Y and, therefore, $\mu_{(X,Y)} \in S_{\Gamma,\mathcal{C},\Upsilon}$.

For the converse, let (X, Y) be such that $\mu_{(X,Y)} \in S_{\Gamma,\mathcal{C},\Upsilon}$. Take $\mathcal{F}_t = \mathcal{F}_t^X \vee \mathcal{F}_t^Y$. Then it follows that on the complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathsf{P})$ the process X_t is adapted to $\mathcal{F}_t$, the process L_t is a $(\mathcal{F}_t)_{t\geq 0}$ adapted symmetric α-stable process and (X, L) satisfies (1.1). Hence (X, L) is indeed a weak solution of (1.1) in the class $\mathcal{V}((1 + \frac{\beta}{\alpha}) \wedge 1)$. $\square$

To complete the proof of Theorem 2.3, we need the following definition.

DEFINITION 5.10. We say that pointwise uniqueness holds in $S_{\Gamma,\mathcal{C},\Upsilon}$ if for any random elements X_1, X_2 and Y defined on the same probability space with $\mu_{(X_1,Y)} \in S_{\Gamma,\mathcal{C},\Upsilon}$ and $\mu_{(X_2,Y)} \in S_{\Gamma,\mathcal{C},\Upsilon}$ one has $X_1 = X_2$ a.s.

Now we are ready to finish.

PROOF OF THEOREM 2.3. By Proposition 2.9, we know that (1.1) has a weak solution in the class $\mathcal{V}((1+\frac{\beta}{\alpha})\wedge 1)$. Thus $\mathcal{S}_{\Gamma,\Upsilon}\neq\varnothing$. By Proposition 2.10 and Proposition 2.8, we know pathwise uniqueness holds for (1.1). Then the converse part of Lemma 5.9 implies that pointwise uniqueness holds in $\mathcal{S}_{\Gamma,\mathcal{C},\Upsilon}$. Then pointwise uniqueness in $\mathcal{S}_{\Gamma,\mathcal{C},\Upsilon}$ along with [31], Lemma 2.10, and the direct part of Lemma 5.9, implies that the hypotheses of [31], Theorem 3.4 are satisfied. Now [31], Theorem 3.4 and Proposition 2.13 together imply that (1.1) has a unique strong solution in the class $\mathcal{V}((1+\frac{\beta}{\alpha})\wedge 1)$. (As a caution to the reader, to avoid any confusion, we note that the word "strong" has a different meaning in [31].) $\square$

Acknowledgments. The authors are very grateful to Khoa Lê for pointing out the correct definition of an approximating sequence in Besov space. The authors also would like to thank Nicolas Perkowski and David Prömel for helpful discussions about Besov spaces and the referees for helpful comments and suggestions. Part of the work on the project has been done during the visits of the authors to the Fields Institute (Toronto, Canada), Technion—Israel Institute of Technology (Haifa, Israel) and Indian Statistical Institute (Bangalore, India). We thank them all for their support and hospitality. OB is very grateful to the Fields institute and especially to Bryan Eelhart for their support, hospitality and incredible coffee breaks. LM is grateful to the Johannes Gutenberg University Mainz where part of this research has been done.

The first author was supported in part by ISF-UGC Grant and CPDA.

The second author was supported in part by ISF-UGC Grant No. 1131/14 and DFG Research Unit FOR 2402.

The third author was supported in part by ISF-UGC Grant No. 1131/14 and Israel Science Foundation grant 1704/18.

SUPPLEMENTARY MATERIAL

Supplement to "Strong existence and uniqueness for stable stochastic differential equations with distributional drift." (DOI: 10.1214/19-AOP1358SUPP; .pdf). The supplementary material recalls some classical definitions and contains proofs of auxiliary results.

REFERENCES

[1] ALDOUS, D. (1978). Stopping times and tightness. *Ann. Probab.* **6** 335–340. MR0474446 https://doi.org/10.1214/aop/1176995579

[2] APPLEBAUM, D. (2009). *Lévy Processes and Stochastic Calculus*, 2nd ed. *Cambridge Studies in Advanced Mathematics* **116**. Cambridge Univ. Press, Cambridge. MR2512800 https://doi.org/10.1017/CBO9780511809781

[3] ATHREYA, S., BUTKOVSKY, O. and MYTNIK, L. (2020). Supplement to "Strong existence and uniqueness for stable stochastic differential equations with distributional drift." https://doi.org/10.1214/19-AOP1358SUPP.

[4] BANDINI, E. and RUSSO, F. (2017). Weak Dirichlet processes with jumps. *Stochastic Process. Appl.* **127** 4139–4189. MR3718109 https://doi.org/10.1016/j.spa.2017.04.001

[5] BAÑOS, D.R., ORTIZ-LATORRE, S., PILIPENKO, A. and PROSKE, F. (2017). Strong solutions of SDE's with generalized drift and multidimensional fractional Brownian initial noise. arXiv preprint, arXiv:1705.01616.

[6] BASS, R. F. and CHEN, Z.-Q. (2001). Stochastic differential equations for Dirichlet processes. *Probab. Theory Related Fields* **121** 422–446. MR1867429 https://doi.org/10.1007/s004400100151

[7] BASS, R. F. and CHEN, Z.-Q. (2003). Brownian motion with singular drift. *Ann. Probab.* **31** 791–817. MR1964949 https://doi.org/10.1214/aop/1048516536

[8] BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd ed. *Wiley Series in Probability and Statistics*: *Probability and Statistics*. Wiley, New York. MR1700749 https://doi.org/10.1002/9780470316962

[9] BLUMENTHAL, R. M. and GETOOR, R. K. (1960). Some theorems on stable processes. *Trans. Amer. Math. Soc.* **95** 263–273. MR0119247 https://doi.org/10.2307/1993291

[10] BOGACHEV, V. I. and PILIPENKO, A. YU. (2016). Strong solutions to stochastic equations with Lévy noise and a discontinuous drift coefficient. *Dokl. Akad. Nauk* **463** 509–513. MR3444003 https://doi.org/10.1134/S1064562415040213

[11] CATELLIER, R. and GUBINELLI, M. (2016). Averaging along irregular curves and regularisation of ODEs. *Stochastic Process. Appl.* **126** 2323–2366. MR3505229 https://doi.org/10.1016/j.spa.2016.02.002

[12] CHEN, Z.-Q., SONG, R. and ZHANG, X. (2018). Stochastic flows for Lévy processes with Hölder drifts. *Rev. Mat. Iberoam.* **34** 1755–1788. MR3896248 https://doi.org/10.4171/rmi/1042

[13] CHEN, Z.-Q. and WANG, L. (2016). Uniqueness of stable processes with drift. *Proc. Amer. Math. Soc.* **144** 2661–2675. MR3477084 https://doi.org/10.1090/proc/12909

[14] COQUET, F., JAKUBOWSKI, A., MÉMIN, J. and SŁOMIŃSKI, L. (2006). Natural decomposition of processes and weak Dirichlet processes. In *In Memoriam Paul-André Meyer: Séminaire de Probabilités XXXIX. Lecture Notes in Math.* **1874** 81–116. Springer, Berlin. MR2276891 https://doi.org/10.1007/978-3-540-35513-7_8

[15] DAVIE, A. M. (2007). Uniqueness of solutions of stochastic differential equations. *Int. Math. Res. Not. IMRN* **24** Art. ID rnm124, 26. MR2377011 https://doi.org/10.1093/imrn/rnm124

[16] ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. Wiley, New York. MR0838085 https://doi.org/10.1002/9780470316658

[17] FLANDOLI, F., GUBINELLI, M. and PRIOLA, E. (2010). Well-posedness of the transport equation by stochastic perturbation. *Invent. Math.* **180** 1–53. MR2593276 https://doi.org/10.1007/s00222-009-0224-4

[18] FLANDOLI, F., ISSOGLIO, E. and RUSSO, F. (2017). Multidimensional stochastic differential equations with distributional drift. *Trans. Amer. Math. Soc.* **369** 1665–1688. MR3581216 https://doi.org/10.1090/tran/6729

[19] FLANDOLI, F., RUSSO, F. and WOLF, J. (2003). Some SDEs with distributional drift. I. General calculus. *Osaka J. Math.* **40** 493–542. MR1988703

[20] FLANDOLI, F., RUSSO, F. and WOLF, J. (2004). Some SDEs with distributional drift. II. Lyons–Zheng structure, Itô's formula and semimartingale characterization. *Random Oper. Stoch. Equ.* **12** 145–184. MR2065168 https://doi.org/10.1163/156939704323074700

[21] FÖLLMER, H. (1981). Dirichlet processes. In *Stochastic Integrals (Proc. Sympos., Univ. Durham, Durham*, 1980). *Lecture Notes in Math.* **851** 476–478. Springer, Berlin. MR0621001

[22] GUBINELLI, M., IMKELLER, P. and PERKOWSKI, N. (2015). Paracontrolled distributions and singular PDEs. *Forum Math. Pi* **3** e6, 75. MR3406823 https://doi.org/10.1017/fmp.2015.2

[23] GUBINELLI, M. and PERKOWSKI, N. (2015). *Lectures on Singular Stochastic PDEs. Ensaios Matemáticos* [*Mathematical Surveys*] **29**. Sociedade Brasileira de Matemática, Rio de Janeiro. MR3445609

[24] HAADEM, S. and PROSKE, F. (2014). On the construction and Malliavin differentiability of solutions of Lévy noise driven SDE's with singular coefficients. *J. Funct. Anal.* **266** 5321–5359. MR3177338 https://doi.org/10.1016/j.jfa.2014.02.009

[25] HAIRER, M. (2014). A theory of regularity structures. *Invent. Math.* **198** 269–504. MR3274562 https://doi.org/10.1007/s00222-014-0505-4

[26] HARRISON, J. M. and SHEPP, L. A. (1981). On skew Brownian motion. *Ann. Probab.* **9** 309–313. MR0606993

[27] JACOD, J. and SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **288**. Springer, Berlin. MR1943877 https://doi.org/10.1007/978-3-662-05265-5

[28] KIM, P. and SONG, R. (2014). Stable process with singular drift. *Stochastic Process. Appl.* **124** 2479–2516. MR3192504 https://doi.org/10.1016/j.spa.2014.03.006

[29] KRYLOV, N. V. and RÖCKNER, M. (2005). Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields* **131** 154–196. MR2117951 https://doi.org/10.1007/s00440-004-0361-z

[30] KULIK, A. M. (2019). On weak uniqueness and distributional properties of a solution to an SDE with α-stable noise. *Stochastic Process. Appl.* **129** 473–506. MR3907007 https://doi.org/10.1016/j.spa.2018.03.010

[31] KURTZ, T. G. (2014). Weak and strong solutions of general stochastic models. *Electron. Commun. Probab.* **19** no. 58, 16. MR3254737 https://doi.org/10.1214/ECP.v19-2833

[32] KURTZ, T. G. and PROTTER, P. (1991). Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.* **19** 1035–1070. MR1112406

[33] KWAŚNICKI, M. (2017). Ten equivalent definitions of the fractional Laplace operator. *Fract. Calc. Appl. Anal.* **20** 7–51. MR3613319 https://doi.org/10.1515/fca-2017-0002

[34] LE GALL, J.-F. (1984). One-dimensional stochastic differential equations involving the local times of the unknown process. In *Stochastic Analysis and Applications* (*Swansea*, 1983). *Lecture Notes in Math*. **1095** 51–82. Springer, Berlin. MR0777514 https://doi.org/10.1007/BFb0099122

[35] LI, Z. and MYTNIK, L. (2011). Strong solutions for stochastic differential equations with jumps. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **47** 1055–1067. MR2884224 https://doi.org/10.1214/10-AIHP389

[36] LI, Z. and PU, F. (2012). Strong solutions of jump-type stochastic equations. *Electron*. *Commun*. *Probab*. **17** no. 33, 13. MR2965746 https://doi.org/10.1214/ECP.v17-1915

[37] PERKOWSKI, N. (2017). Paracontrolled distributions and singular diffusions. Preprint, http://www.mathematik.hu-berlin.de/~perkowsk/files/teaching/paracontrolled-bonn.pdf.

[38] PORTENKO, N. I. (1979). Diffusion processes with a generalized drift coefficient. *Teor*. *Veroyatn*. *Primen*. **24** 62–77. MR0522237

[39] PORTENKO, N. I. (1990). *Generalized Diffusion Processes*. *Translations of Mathematical Monographs* **83**. Amer. Math. Soc., Providence, RI. Translated from the Russian by H. H. McFaden. MR1104660

[40] PRIOLA, E. (2012). Pathwise uniqueness for singular SDEs driven by stable processes. *Osaka J*. *Math*. **49** 421–447. MR2945756

[41] PRIOLA, E. (2018). Davie's type uniqueness for a class of SDEs with jumps. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **54** 694–725. MR3795063 https://doi.org/10.1214/16-AIHP818

[42] TANAKA, H., TSUCHIYA, M. and WATANABE, S. (1974). Perturbation of drift-type for Lévy processes. *J*. *Math*. *Kyoto Univ*. **14** 73–92. MR0368146 https://doi.org/10.1215/kjm/1250523280

[43] TRIEBEL, H. (2010). *Theory of Function Spaces*. *Modern Birkhäuser Classics*. Birkhäuser/Springer Basel AG, Basel. Reprint of 1983 edition [MR0730762], Also published in 1983 by Birkhäuser Verlag [MR0781540]. MR3024598

[44] VERETENNIKOV, A. JU. (1980). Strong solutions and explicit formulas for solutions of stochastic integral equations. *Mat*. *Sb*. **111(153)** 434–452, 480. MR0568986

[45] ZHANG, X. and ZHAO, G. (2018). Heat kernel and ergodicity of SDEs with distributional drifts. arXiv preprint, arXiv:1710.10537.

[46] ZVONKIN, A. K. (1974). A transformation of the phase space of a diffusion process that will remove the drift. *Mat*. *Sb*. **93(135)** 129–149, 152. MR0336813