STRONG CONVERGENCE OF THE EULER–MARUYAMA APPROXIMATION FOR A CLASS OF LÉVY-DRIVEN SDES

FRANZISKA KÜHN AND RENÉ L. SCHILLING

Abstract. Consider the following stochastic differential equation (SDE)

$$dX_t = b(t, X_{t-})\,dt + dL_t, \quad X_0 = x,$$

driven by a d-dimensional Lévy process $(L_t)_{t\geq 0}$. We establish conditions on the Lévy process and the drift coefficient b such that the Euler–Maruyama approximation converges strongly to a solution of the SDE with an explicitly given rate. The convergence rate depends on the regularity of b and the behaviour of the Lévy measure at the origin. As a by-product of the proof, we obtain that the SDE has a pathwise unique solution. Our result covers many important examples of Lévy processes, e.g. isotropic stable, relativistic stable, tempered stable and layered stable.

1. Introduction

For a given Lévy process $(L_t)_{t\geq 0}$ with values in $\mathbb{R}^d$ and Lévy triplet (ℓ, Q, ν) we consider the stochastic differential equation (SDE)

$$dX_t = b(t, X_{t-})\,dt + dL_t, \quad X_0 = x \in \mathbb{R}^d. \tag{1}$$

If the drift coefficient b is Hölder continuous in time and space, there is a quite general result on the existence of a pathwise unique solution, cf. Chen, Song & Zhang [1]. It is, however, in general not possible to calculate the solution explicitly, and therefore it is important to have numerical schemes which allow us to approximate the solution. In this paper we derive conditions on the Lévy process $(L_t)_{t\geq 0}$ and the drift coefficient b such that the Euler–Maruyama approximation

$$X_t^{(n)} - x = \int_0^t b\big(s, X^{(n)}_{\eta_n(s)-}\big) + L_t, \quad t \in [0, T],\ n \in \mathbb{N},$$

converges strongly to a solution of the given SDE with a certain rate; here $\eta_n(s) := T\frac{i}{n}$ for $s \in [T\frac{i}{n}, T\frac{i+1}{n})$. It turns out that the convergence (rate) depends on two factors: the regularity of b and the behaviour of the Lévy measure at the origin.

If $b = b(x)$ satisfies a one-sided Lipschitz condition, then a result by Higham & Kloeden [5] shows that the Euler–Maruyama approximation converges strongly with convergence rate 1/2. It is natural to ask whether the regularity assumption can be weakened to Hölder regularity. Pamen & Taguchi [20] study the convergence rate for SDEs with Hölder continuous coefficients driven by Brownian motion and by truncated α-stable Lévy processes with index $\alpha > 1$. For isotropic α-stable Lévy processes, $\alpha > 1$, first results were obtained by Qiao [22] (b is Lipschitz up to a log-term) and by Hashimoto [4] who proves the strong convergence under a Komatsu condition, but does not determine the convergence rate. More recently, Mikulevicius & Xu [18] have shown that

$$\mathbb{E}\left(\sup_{0\leq t\leq T} |X_t - X_t^{(n)}|^p\right) \leq Cn^{-p\beta/\alpha}, \quad p \in (0, \alpha),$$

if $\alpha \in (1, 2)$ and b is β-Hölder continuous for some $\beta > 1 - \alpha/2$. This estimate shows, in particular, that there are two factors which result in a slow convergence of the Euler–Maruyama approximation: weak regularity of b and a strong singularity of the Lévy measure $\nu(dy) = |y|^{-(d+\alpha)}\,dy$ at $y = 0$. The fact that the behaviour of the Lévy measure influences the convergence rate was already observed by Jacod [7] who investigated the weak convergence of the Euler–Maruyama approximation for a class of Lévy-driven SDEs.

2010 *Mathematics Subject Classification.* Primary: 60H10. Secondary: 60G51, 60J35, 60J75.

Key words and phrases. Euler–Maruyama approximation, stochastic differential equation, strong convergence.

Our main result, Theorem 2.1, shows the strong convergence of the Euler–Maruyama approximation for a large class of driving Lévy processes covering many important and interesting examples, e.g. isotropic α-stable, relativistic stable, tempered stable and layered stable Lévy processes. The proof relies on the so-called Itô-Tanaka trick which relates the time average $\int_0^t b(s, X_s)\,ds$ of the solution $(X_t)_{t\geq 0}$ to (1) with the solution u to the Kolmogorov equation

$$\partial_t u(t,x) + A_x u(t,x) + b(t,x)\cdot\nabla_x u(t,x) = -b(t,x); \tag{2}$$

here A_x denotes the infinitesimal generator of the driving Lévy process $(L_t)_{t\geq 0}$ acting with respect to the space variable x. The key step is to prove the existence of a solution to (2) which is sufficiently regular and satisfies certain Hölder estimates. The required regularity of u depends on the regularity of b and the behaviour of the Lévy measure at 0.

The Itô-Tanaka trick has been used by Pamen & Taguchi [20] to prove the strong convergence of the Euler–Maruyama approximation for the particular case that $(L_t)_{t\geq 0}$ is a Brownian motion or a truncated stable Lévy process taking values in $\mathbb{R}^d$, $d\geq 2$. For these two processes the existence of a sufficiently nice solution to the Kolmogorov equation (2) was already known. Pamen & Taguchi do not take advantage of the fact that the required regularity of the solution depends on the behaviour of the Lévy measure at 0, and therefore they end up with a convergence rate which is far from being optimal.

Our paper is organized as follows. In Section 2 we state and discuss the main results; the required definitions will be explained in Section 3. Section 4 is devoted to the proofs of the main results, and in Section 5 we illustrate our results with examples. Some auxiliary statements are proved in the appendix.

2. Main results

2.1. **Theorem.** *Let $(L_t)_{t\geq 0}$ be a d-dimensional Lévy process with Lévy triplet $(\ell, 0, \nu)$ and characteristic exponent $\psi : \mathbb{R}^d \to \mathbb{C}$, and let $\gamma_0 \in [1,2]$, $\gamma_\infty > 0$ be such that $\int_{|z|\leq 1} |z|^{\gamma_0}\,\nu(dz) < \infty$ and $\int_{|z|\geq 1} |z|^{\gamma_\infty}\,\nu(dz) < \infty$. Assume that L_t admits a transition density $p_t \in C^2(\mathbb{R}^d)$ for all $t > 0$, such that there exist constants $\alpha \in (1,2]$ and $c = c(T) > 0$ such that*

$$\int_{\mathbb{R}^d} |\partial_{x_i} p_t(x)|\,dx \leq c t^{-1/\alpha} \quad \text{for all} \quad i \in \{1,\dots,d\},\ t\in(0,T]. \tag{3}$$

Let $b : [0,\infty)\times\mathbb{R}^d \to \mathbb{R}^d$ be a bounded function which is β-Hölder continuous with respect to x and η-Hölder continuous with respect to t for some $\beta,\eta \in (0,1]$, i.e.

$$|b(t,x) - b(t,y)| \leq C|x-y|^\beta \quad \text{and} \quad |b(s,x) - b(t,x)| \leq C|s-t|^\eta$$

holds for all $s,t \geq 0$, $x,y\in\mathbb{R}^d$ and with an absolute constant $C > 0$. If the balance condition

$$2\alpha - \gamma_0(1-\beta) > 2, \tag{4}$$

is satisfied, then the SDE

$$dX_t = b(t, X_{t-})\,dt + dL_t, \quad X_0 = x \tag{5}$$

has a pathwise unique strong solution $(X_t)_{t\geq 0}$, and for any $p \leq \gamma_\infty$ and $T > 0$ there exists a constant $C > 0$ such that

$$\mathbb{E}\left(\sup_{0\leq t\leq T} |X_t - X_t^{(n)}|^p\right) \leq C n^{-\min\{1,\, p\beta/\gamma_0,\, p\eta\}} \quad \text{for all} \quad n\in\mathbb{N}. \tag{6}$$

2.2. **Remark.** (i) For the existence of a pathwise unique solution to (5) it is crucial that the mapping $x \mapsto b(t,x)$ is sufficiently regular. For instance, if $(L_t)_{t\geq 0}$ is an isotropic α-stable Lévy process, then the SDE

$$dX_t = b(X_{t-})\,dt + dL_t$$

fails, in general, to have a pathwise unique solution if b is β-Hölder continuous with $\beta+\alpha < 1$, cf. [27]; recently, Kulik [15] has shown that the SDE admits a pathwise unique solution if $\beta + \alpha > 1$. This shows that there is a trade-off and compensation between the (lack of) regularity of the driving Lévy noise and the (lack of) regularity of the coefficient $x\mapsto b(x)$.

(ii) Condition (3) is equivalent to saying that the semigroup $P_t\phi(x) := \mathbb{E}\phi(x+L_t)$, $\phi \in \mathcal{B}_b(\mathbb{R}^d)$, associated with the Lévy process $(L_t)_{t\geq 0}$ satisfies the gradient estimate

$$\|\nabla P_t \phi\|_\infty \leq c t^{-1/\alpha} \|\phi\|_\infty, \quad \phi \in \mathcal{B}_b(\mathbb{R}^d),$$

cf. Lemma 4.1. If $(L_t)_{t\geq 0}$ is subordinate to a Brownian motion, then (3) can also be understood as a moment estimate, cf. Lemma 4.5.

(iii) The existence of the moments $\int_{|z|\leq 1} |z|^{\gamma_0}\,\nu(dz)$ and $\int_{|z|\geq 1} |z|^{\gamma_\infty}\,\nu(dz)$ is related to the growth of the characteristic exponent ψ, cf. Lemma 5.1; Lemma 5.1 is very useful since it allows us to verify the assumptions of Theorem 2.1 if the Lévy measure ν cannot be calculated explicitly.

(iv) A sufficient condition for the existence of a transition probability density $p_t \in C^2(\mathbb{R}^d)$ for all $t > 0$ is the Hartman–Wintner condition:

$$\lim_{|\xi|\to\infty} \frac{\operatorname{Re}\psi(\xi)}{\log(1+|\xi|)} = \infty,$$

cf. [8] for a thorough discussion.

(v) Let $(X_t)_{t\geq 0}$ be a solution to (5). Since b is bounded, we have for any $t > 0$

$$X_t \in L^p(\mathbb{P}) \iff L_t \in L^p(\mathbb{P}) \iff \int_{|z|\geq 1} |z|^p\,\nu(dz) < \infty$$

– for the second equivalence see Sato [23] –, i.e. the solution inherits the integrability of the driving Lévy process and vice versa. This means that, in general, we cannot expect (6) to hold for $p > \gamma_\infty$.

(vi) A slight variation of our arguments, see the uniqueness part of the proof of Theorem 2.1 on page 15, allows us to derive the estimate

$$\mathbb{E}\left(\sup_{t\in[0,T]} |X_t(x) - X_t(y)|^p\right) \leq C'|x-y|^p, \qquad p \leq \gamma_\infty$$

for some constant $C' > 0$ where $X_t(z)$ denotes the solution to the SDE with initial condition $X_0(z) = z$. If $\gamma_\infty > d$ it follows from a standard Kolmogorov–Chentsov–Totoki argument that $x \mapsto X_t(x)$ is Hölder continuous of order $\kappa < 1 - d/\gamma_\infty$.

Let us give some further remarks on possible extensions of Theorem 2.1.

2.3. **Remark.** (i) If $(L_t)_{t\geq 0}$ has a non-vanishing (possibly degenerate) diffusion part, then the statement of Theorem 2.1 remains valid for $\gamma_0 := 2$. In particular, if $(W_t)_{t\geq 0}$ is a Brownian motion and $b : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$ is a bounded function which is β-Hölder continuous with respect to x and η-Hölder-continuous with respect to t for some $\beta, \eta \in (0,1]$, then the SDE

$$dX_t = b(t, X_t)\,dt + dW_t, \quad X_0 = x$$

has a pathwise unique solution, and for any $p > 0$, $T > 0$ there exists a constant $C > 0$ such that

$$\mathbb{E}\left(\sup_{0\leq t\leq T} |X_t - X_t^{(n)}|^p\right) \leq C n^{-\min\{1,\, p\beta/2,\, p\eta\}} \quad \text{for all} \quad n \in \mathbb{N};$$

this result extends [20, Theorem 2.11].

(ii) A close inspection of our arguments reveals that the Hölder condition on $t \mapsto b(t,x)$ can be replaced by uniform continuity; if we denote by

$$w(\delta) := \sup_{x\in\mathbb{R}^d} \sup_{|s-t|\leq\delta} |b(t,x) - b(s,x)|$$

the modulus of continuity of $t \mapsto b(t,x)$ (uniformly in x), then (6) becomes

$$\mathbb{E}\left(\sup_{0\leq t\leq T} |X_t - X_t^{(n)}|^p\right) \leq C n^{-\min\{1, p\beta/\gamma_0\}} + C w(1/n)^p, \qquad n \in \mathbb{N},$$

for a suitable constant $C > 0$.

(iii) Case $\alpha = 1$: If we replace (3) by

$$\int_{\mathbb{R}^d} |\partial_{x_i} p_t(x)|\,dx \le ct^{-1}\log^{-1-\epsilon} t, \quad i \in \{1,\dots,d\},\ t \in (0,T]$$

for some $\epsilon > 0$, then the statement of Theorem 2.1 holds for $\alpha = 1$.

Combining Theorem 2.1 with the gradient estimates in [26] we can easily prove the following statement which covers many interesting and important examples of Lévy processes, cf. Section 5.

2.4. **Corollary.** *Let $(L_t)_{t\ge 0}$ be a d-dimensional Lévy process with characteristic exponent ψ and Lévy triplet $(\ell, 0, \nu)$. Let $\gamma_0 \in [1,2]$, $\gamma_\infty > 0$ be exponents such that $\int_{|z|\le 1} |z|^{\gamma_0}\,\nu(dz) < \infty$ and $\int_{|z|\ge 1} |z|^{\gamma_\infty}\,\nu(dz) < \infty$. Assume that there exists a strictly increasing function $f : (0,\infty) \to [0,\infty)$ which is differentiable near infinity and satisfies the following conditions.*

(i) $c^{-1} f(|\xi|) \le \operatorname{Re}\psi(\xi) \le cf(|\xi|)$ *as* $|\xi| \to \infty$ *for some constant* $c \in (0,\infty)$*;*
(ii) $\limsup_{r\to\infty} f^{-1}(2r)/f^{-1}(r) < \infty$*;*
(iii) *there exist constants* $\alpha \in (1,2]$ *and* $c > 0$ *such that* $f(r) \ge cr^\alpha$ *for large* $r > 0$*.*

Let $b : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$ be a bounded function which is β-Hölder continuous with respect to x and η-Hölder continuous with respect to t for some $\beta, \eta \in (0,1]$. If the balance condition holds

$$2\alpha - \gamma_0(1-\beta) > 2,$$

then the SDE

$$dX_t = b(t, X_{t-})\,dt + dL_t, \quad X_0 = x$$

has a pathwise unique strong solution $(X_t)_{t\ge 0}$, and for any $p \le \gamma_\infty$ and $T > 0$ there exists a constant $C > 0$ such that

$$\mathbb{E}\left(\sup_{0\le t\le T} |X_t - X_t^{(n)}|^p\right) \le Cn^{-\min\{1,\,p\beta/\gamma_0,\,p\eta\}} \quad \textit{for all} \ \ n \in \mathbb{N}.$$

Typical examples for f are $f(r) = r^\alpha$ and $f(r) = r^\alpha \log^\beta(1+r)$, see Section 5.

For the particular case that the driving Lévy process is subordinate to a Brownian motion, Theorem 2.1 has the following corollary.

2.5. **Corollary.** *Let $L_t = B_{S_t}$ be a d-dimensional Brownian motion subordinated by a subordinator $(S_t)_{t\ge 0}$ with Laplace exponent (Bernstein function) f,*

$$f(\lambda) = \int_{(0,\infty)} (1 - e^{-\lambda r})\,\mu(dr), \quad \lambda \ge 0.$$

Assume that there exist constants $\delta_0 \in [1/2, 1]$, $\delta_\infty > 0$ and $\rho \in (1/2, 1]$ such that

$$\int_{(0,1)} r^{\delta_0}\,\mu(dr) + \int_{(1,\infty)} r^{\delta_\infty}\,\mu(dr) < \infty$$

and

$$\liminf_{\lambda\to\infty} \frac{f(\lambda)}{\lambda^\rho} > 0.$$

Let $b : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$ be a bounded function which is β-Hölder continuous with respect to x and η-Hölder continuous with respect to t for some $\eta \in (0,1]$. If

$$2\rho - \delta_0(1-\beta) > 1$$

then the SDE

$$dX_t = b(t, X_{t-})\,dt + dL_t, \quad X_0 = x$$

has a unique strong solution $(X_t)_{t\ge 0}$ and for any $p \le 2\delta_\infty$ and $T > 0$ there exists a constant $C > 0$ such that

$$\mathbb{E}\left(\sup_{0\le t\le T} |X_t - X_t^{(n)}|^p\right) \le Cn^{-\min\{1,\,p\beta/(2\delta_0),\,p\eta\}} \quad \textit{for all} \ \ n \in \mathbb{N}.$$

3. Preliminaries

We consider Euclidean space $\mathbb{R}^d$ endowed with the Borel σ-algebra $\mathcal{B}(\mathbb{R}^d)$. The open ball centered at $x \in \mathbb{R}^d$ of radius $r > 0$ is denoted by $B(x,r)$. For a differentiable function $f : \mathbb{R}^d \to \mathbb{R}$ the partial derivative with respect to x_i is denoted by $\partial_{x_i} f$, and ∇f is the gradient of f. As usual, $C^k(\mathbb{R}^d)$ is the space of k-times continuously differentiable functions, $\mathcal{B}_b(\mathbb{R}^d)$ the space of bounded Borel measurable functions, and $C_\infty(\mathbb{R}^d)$ is the space of continuous functions vanishing at infinity. For $\beta \in [0,1]$ we define Hölder spaces by

$$\mathcal{C}_b^\beta(\mathbb{R}^d) := \left\{ f : \mathbb{R}^d \to \mathbb{R}^k;\ \|f\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} := \sup_{x \in \mathbb{R}^d} |f(x)| + \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x-y|^\beta} < \infty \right\}$$

$$\mathcal{C}_b^{1,\beta}(\mathbb{R}^d) := \left\{ f \in C^1(\mathbb{R}^d, \mathbb{R}^k);\ \|f\|_{\mathcal{C}_b^{1,\beta}(\mathbb{R}^d)} := \sup_{x \in \mathbb{R}^d} |f(x)| + \|\nabla f\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} < \infty \right\}.$$

For a function space M and a function $g : [a,b] \times \mathbb{R}^d \to \mathbb{R}^k$ we write $g \in C([a,b], M)$ if $g(t,\cdot) \in M$ for all $t \in [a,b]$ and $t \mapsto g(t,\cdot)$ is continuous. Similarly, $g \in C^1([a,b], M)$ means that $\partial_t g(t,\cdot) \in M$ and $t \mapsto \partial_t g(t,\cdot)$ is continuous. If M is a normed function space, then

$$\|g\|_{C([a,b],M)} := \sup_{t \in [a,b]} \|g(t,\cdot)\|_M$$

defines a norm on $C([a,b], M)$. For brevity we will often denote this norm by $\|g\|_M$; in particular we will write

$$\|g\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} := \sup_{t \in [a,b]} \sup_{x \in \mathbb{R}^d} |g(t,x)| + \sup_{t \in [a,b]} \sup_{x \neq y} \frac{|g(t,x) - g(t,y)|}{|x-y|^\beta}.$$

We say that a bounded function $g : [a,b] \times \mathbb{R}^d \to \mathbb{R}^k$ is *β-Hölder continuous with respect to x* if $\|g\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} < \infty$.

Throughout, $(\Omega, \mathcal{A}, \mathbb{P})$ is a probability space. A family of random variables $L_t : \Omega \to \mathbb{R}^d$, $t \geq 0$, is a *d-dimensional Lévy process* if $(L_t)_{t \geq 0}$ has stationary and independent increments, $t \mapsto L_t$ is, with probability 1, right-continuous with finite left limits (càdlàg), and $L_0 = 0$. A Lévy process can be uniquely (in the sense of finite-dimensional distributions) characterized by its *characteristic exponent* $\psi : \mathbb{R}^d \to \mathbb{C}$,

$$\mathbb{E} e^{i\xi \cdot L_t} = e^{-t\psi(\xi)}, \quad t \geq 0,\ \xi \in \mathbb{R}^d;$$

the exponent is given by the *Lévy–Khintchine formula*

$$\psi(\xi) = -i\ell \cdot \xi + \frac{1}{2} \xi \cdot Q\xi + \int_{y \neq 0} \left(1 - e^{iy\cdot\xi} + iy \cdot \xi \mathbb{1}_{(0,1)}(|y|)\right) \nu(dy), \quad \xi \in \mathbb{R}^d.$$

There is a one-to-one correspondence between the exponent ψ and the *Lévy triplet* (ℓ, Q, ν) consisting of a vector $\ell \in \mathbb{R}^d$ (*drift parameter*), a symmetric positive semi-definite matrix $Q \in \mathbb{R}^{d \times d}$ (*diffusion parameter*) and a measure ν on $(\mathbb{R}^d \setminus \{0\}, \mathcal{B}(\mathbb{R}^d \setminus \{0\}))$ satisfying $\int_{y \neq 0} \min\{|y|^2, 1\}\, \nu(dy) < \infty$ (*Lévy measure*).

It is not difficult to see that any Lévy process $(L_t)_{t \geq 0}$ is a Markov process, and therefore there is a transition semigroup $(P_t)_{t \geq 0}$ and an infinitesimal generator $(A, \mathcal{D}(A))$ associated with $(L_t)_{t \geq 0}$. It is well known that

$$P_t f(x) = \mathbb{E} f(x + L_t), \quad x \in \mathbb{R}^d,\ t \geq 0,\ f \in \mathcal{B}_b(\mathbb{R}^d),$$

and, for any $f \in C_b^2(\mathbb{R}^d)$,

$$Af(x) = \ell \cdot \nabla f(x) + \frac{1}{2} \operatorname{div}\left(Q \nabla f(x)\right) + \int_{y \neq 0} \left(f(x+y) - f(x) - \nabla f(x) \cdot y \mathbb{1}_{(0,1)}(|y|)\right) \nu(dy). \tag{7}$$

If $(L_t)_{t \geq 0}$ is a Lévy process with Lévy triplet $(\ell, 0, \nu)$ and $\int_{|y| \leq 1} |y|^\gamma\, \nu(dy) < \infty$ for some $\gamma \in [1,2]$, then $\mathcal{C}_b^{1,\gamma-1}(\mathbb{R}^d) \subseteq \mathcal{D}(A)$, and (7) holds for any $f \in \mathcal{C}_b^{1,\gamma-1}(\mathbb{R}^d)$, cf. [13, Theorem 4.1]; moreover,

$$\|Af\|_\infty \leq 2 \left(|\ell| + \int_{y \neq 0} \min\{|y|^\gamma, 1\}\, \nu(dy) \right) \|f\|_{\mathcal{C}_b^{1,\gamma-1}(\mathbb{R}^d)} \quad \text{for all} \quad f \in \mathcal{C}_b^{1,\gamma-1}(\mathbb{R}^d). \tag{8}$$

For a function $f = f(t,x)$ we indicate by $A_x f(t,x)$ that the operator A acts on the variable x for any fixed t. Our standard reference for Lévy processes is the monograph [23] by Sato.

A Lévy process $(S_t)_{t\geq 0}$ with non-decreasing sample paths is called a *subordinator*. It can be uniquely characterized by its Laplace exponent (Bernstein function) $f(r) := \log \mathbb{E}e^{-rS_t}$,

$$f(r) = br + \int_{(0,\infty)} (1 - e^{\lambda r})\, \mu(dr), \quad r > 0,$$

where $b \geq 0$ and μ is a measure on $(0,\infty)$ such that $\int_{(0,\infty)} \min\{r,1\}\, \mu(dr) < \infty$. If $(L_t)_{t\geq 0}$ is a Lévy process with characteristic exponent ψ and $(S_t)_{t\geq 0}$ a subordinator with Laplace exponent f such that $(L_t)_{t\geq 0}$ and $(S_t)_{t\geq 0}$ are independent, then the subordinate process

$$X_t(\omega) := L_{S_t(\omega)}(\omega), \quad \omega \in \Omega,\ t \geq 0,$$

is a Lévy process and its characteristic exponent is given by $f(\psi(\xi))$, see [25] for further details.

A stochastic differential equation (SDE) driven by a Lévy process is of the form

$$dX_t = b(t, X_{t-})\, dt + g(t, X_{t-})\, dL_t$$

for suitable coefficients b, g and an initial condition fixing X_0. The *Euler–Maruyama approximation* of the SDE is given by

$$X_t^{(n)} := X_0 + \int_0^t b(s, X^{(n)}_{\eta_n(s)-})\, ds + \int_0^t g(s, X^{(n)}_{\eta_n(s)-})\, dL_s, \quad t \in [0,T] \tag{9}$$

where $\eta_n(s) := T\frac{i}{n}$ for any $s \in [T\frac{i}{n}, T\frac{i+1}{n})$, $i = 0, 1, \ldots, n$, for fixed $T > 0$. We say that the SDE has a *pathwise unique* solution if for any two solutions $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ such that $X_0 = Y_0$, we have

$$\mathbb{P}\,(\forall t \geq 0 : X_t = Y_t) = 1.$$

We refer to Ikeda–Watanabe [6] and Protter [21] for a thorough discussion of stochastic integration and SDEs.

4. Proofs

Before we start to prove Theorem 2.1 let us briefly explain the idea of the proof. Suppose that $(X_t)_{t\geq 0}$ solves the SDE

$$dX_t = b(t, X_{t-})\, dt + dL_t, \quad X_0 = x.$$

From the definition of the Euler–Maruyama approximation (9) we see that

$$X_t - X_t^{(n)} = \int_0^t \Big(b(s, X_{s-}) - b(s, X^{(n)}_{\eta_n(s)-})\Big)\, ds, \tag{10}$$

and so we have to show that the right-hand side converges in $L^p(\mathbb{P})$ to 0 as $n \to \infty$. To this end, we use the so-called Itô-Tanaka trick: We will show that there exists a sufficiently well-behaved solution $u : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ to the integro-differential equation

$$\frac{\partial}{\partial t} u(t,x) + A_x u(t,x) + b(t,x) \cdot \nabla_x u(t,x) = -b(t,x), \quad u(T,x) = 0 \tag{11}$$

for small $T > 0$; by A we denote the generator of the driving Lévy process $(L_t)_{t\geq 0}$. Applying Itô's formula, we get

$$\int_0^t (b(s, X_{s-}) - b(s, X^{(n)}_{\eta_n(s)-})\, ds \approx u(t, X_t) - u(t, X_t^{(n)}) + M_t$$

for some martingale M. If u is sufficiently smooth, this will allow us to estimate the L^p-norm of right-hand side of (10) using the Burkholder–Davis–Gundy inequality, see pp. 11.

In the first part of this section we establish the existence of a solution to (11), cf. Theorem 4.4. In order to make sense of (11) we have, in particular, to show that $u(\cdot, x)$ and $u(t, \cdot)$ are differentiable and that $A_x u(t,x)$ is well-defined. We start with an auxiliary result showing that (3) gives automatically an estimate for the integrated second derivatives $\int_{\mathbb{R}^d} |\partial_{x_i}\partial_{x_i} p_t(x)|\, dx$.

4.1. **Lemma.** *Let $(L_t)_{t\geq 0}$ be a d-dimensional Lévy process with transition semigroup $(P_t)_{t\geq 0}$ and density $p_t \in C^2(\mathbb{R}^d)$. Let $c : (0,T] \to [0,\infty)$ be a non-negative function.*

(i) $\int_{\mathbb{R}^d} |\partial_{x_i} p_t(x)|\, dx \leq c(t) \iff \forall \phi \in \mathcal{B}_b(\mathbb{R}^d) : \|\partial_{x_i} P_t \phi\| \leq c(t)\|\phi\|_\infty$. *If one (hence both) of the conditions is satisfied, then $\nabla P_t \phi = P_t(\nabla \phi)$ for any $\phi \in C_b^1(\mathbb{R}^d)$.*

(ii) *If* $\|\partial_{x_i} P_t\phi\|_\infty \leq c(t)\|\phi\|_\infty$ *for all* $t \in (0,T]$, $\phi \in \mathcal{B}_b(\mathbb{R}^d)$ *and* $i \in \{1,\dots,d\}$, *then* $\|\partial_{x_i}\partial_{x_k} P_{2t}\phi\|_\infty \leq c(t)^2\|\phi\|_\infty$ *for all* $t \in (0,T]$, $\phi \in \mathcal{B}_b(\mathbb{R}^d)$ *and* $i,k \in \{1,\dots,d\}$.

(iii) *If* $\int_{\mathbb{R}^d} |\partial_{x_i} p_t(x)|\,dx \leq c(t)$ *for all* $t \in (0,T]$ *and* $i \in \{1,\dots,d\}$, *then* $\int_{\mathbb{R}^d} |\partial_{x_i}\partial_{x_k} p_{2t}(x)|\,dx \leq c(t)^2$ *for all* $t \in (0,T]$ *and* $i,k \in \{1,\dots,d\}$.

Proof. (i) Suppose that $\int_{\mathbb{R}^d} |\partial_{x_i} p_t(x)|\,dx \leq c(t)$ for some $t \in (0,T]$. Since

$$x \mapsto \int_{\mathbb{R}^d} \phi(z)\partial_{x_i} p_t(z-x)\,dz$$

is continuous, it follows from the differentiation lemma, cf. Lemma A.1, that

$$x \mapsto P_t\phi(x) = \mathbb{E}\phi(x+L_t) = \int_{\mathbb{R}^d} \phi(z)p_t(z-x)\,dz$$

is differentiable and

$$\partial_{x_i} P_t\phi(x) = \int_{\mathbb{R}^d} \phi(z)\partial_{x_i} p_t(z-x)\,dz; \tag{12}$$

thus

$$\|\partial_{x_i} P_t\phi\|_\infty \leq c(t)\|\phi\|_\infty, \quad \phi \in \mathcal{B}_b(\mathbb{R}^d). \tag{$\star$}$$

Suppose that ($\star$) holds. Let $\chi_n \in C_c(\mathbb{R}^d)$ be a cut-off function such that $\mathbb{1}_{B(0,n)} \leq \chi_n \leq \mathbb{1}_{B(0,n+1)}$. Applying the differentiation lemma, we find that

$$\partial_{x_i} P_t(\phi\chi_n)(x) = \int_{\mathbb{R}^d} \phi(z)\chi_n(z)\partial_{x_i} p_t(z-x)\,dz = -\int_{\mathbb{R}^d} \phi(x+y)\chi_n(x+y)\partial_{y_i} p_t(y)\,dy$$

for any $x \in \mathbb{R}^d$ and $\phi \in \mathcal{B}_b(\mathbb{R}^d)$. Thus,

$$\begin{aligned}
\int_{|y|\leq n} |\partial_{y_i} p_t(y)|\,dy &= \sup\left\{\left|\int_{|y|\leq n} \phi(y)\partial_{y_i} p_t(y)\,dy\right|;\ \phi \in \mathcal{B}_b(\mathbb{R}^d),\ \|\phi\|_\infty \leq 1\right\} \\
&= \sup\left\{\left|\int_{|y|\leq n} \phi(y)\chi_n(y)\partial_{y_i} p_t(y)\,dy\right|;\ \phi \in \mathcal{B}_b(\mathbb{R}^d),\ \|\phi\|_\infty \leq 1\right\} \\
&\leq \sup\left\{\|\partial_{x_i} P_t(\phi\chi_n)\|_\infty;\ \phi \in \mathcal{B}_b(\mathbb{R}^d),\ \|\phi\|_\infty \leq 1\right\} \leq c(t).
\end{aligned}$$

As $n \in \mathbb{N}$ is arbitrary, the monotone convergence theorem gives $\int_{\mathbb{R}^d} |\partial_{y_i} p_t(y)|\,dy \leq c(t)$. Then (12) and the integration by parts formula show that $\nabla P_t\phi = P_t(\nabla\phi)$ for any $\phi \in C_b^1(\mathbb{R}^d)$.

(ii) Fix $\phi \in \mathcal{B}_b(\mathbb{R}^d)$ and $t \in (0,T]$. Since $P_t\phi \in C_b^1(\mathbb{R}^d)$ it follows from (i) and the semigroup property that

$$\begin{aligned}
\|\partial_{x_i}\partial_{x_k} P_{2t}\phi\|_\infty = \|\partial_{x_i}\partial_{x_k} P_t P_t\phi\|_\infty = \|\partial_{x_i} P_t(\partial_{x_k} P_t\phi)\|_\infty &\leq c(t)\|\partial_{x_k} P_t\phi\|_\infty \\
&\leq c(t)^2\|\phi\|_\infty.
\end{aligned}$$

(iii) By (i) and (ii), we have $\|\partial_{x_i}\partial_{x_k} P_{2t}\phi\|_\infty \leq c(t)^2\|\phi\|_\infty$. Using a very similar reasoning as in the proof of (i), we find that

$$\int_{|y|\leq n} |\partial_{y_i}\partial_{y_k} p_{2t}(y)|\,dy \leq c(t)^2;$$

applying the monotone convergence theorem completes the proof. □

Recall that we use for a function $g : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ and $\beta \in (0,1]$ the notation

$$\|g\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} := \sup_{t\in[0,T]} \sup_{x\in\mathbb{R}^d} |g(t,x)| + \sup_{t\in[0,T]} \sup_{x\neq y} \frac{|g(t,x)-g(t,y)|}{|x-y|^\beta}.$$

4.2. **Lemma.** *Let $(L_t)_{t\geq 0}$ be a Lévy process as in Theorem 2.1 with generator $(A, \mathcal{D}(A))$. Let $g \in C([0,T], \mathcal{C}_b^\beta(\mathbb{R}^d))$ for $\beta \in (0,1]$ satisfying* (4)*. For every $T > 0$ there exists a mapping $u \in C([0,T], \mathcal{C}_b^{1,\max\{\beta,\gamma_0/2\}}(\mathbb{R}^d)) \cap C^1([0,T], C_b(\mathbb{R}^d))$ solving*

$$\frac{\partial}{\partial t} u(t,x) = A_x u(t,x) + g(t,x) \quad \text{on } [0,T]\times\mathbb{R}^d \tag{13}$$

such that $u(0,\cdot) = 0$ and

$$\|u\|_\infty + \|\nabla u\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} + \|\nabla u\|_{\mathcal{C}_b^{\gamma_0/2}(\mathbb{R}^d)} \leq C(T)\|g\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} \tag{14}$$

for some constant $C(T) > 0$ which does not depend on g and satisfies $\lim_{T\to 0} C(T) = 0$.

4.3. **Remark.** (i) If we define $v(t,x) := u(T-t,x)$ for fixed $T > 0$, then v is a solution to the equation with reversed time

$$-\frac{\partial}{\partial t} v(t,x) = A_x v(t,x) + g(t,x), \quad v(T,\cdot) = 0.$$

(ii) In Theorem 2.1 (and hence in Lemma 4.2) we assume that the constant α appearing in (3) is strictly larger than 1 and we require that the balance condition (4) is satisfied. Both assumptions are crucial for the proof of Lemma 4.2. The assumption $\alpha > 1$ is needed to prove that $\nabla u(t,\cdot)$ exists and $\nabla u(t,\cdot) \in \mathcal{C}_b^\beta(\mathbb{R}^d)$ whereas (4) is used to show that $\nabla u(t,\cdot) \in \mathcal{C}_b^{\gamma_0/2}(\mathbb{R}^d)$. The fact that $\nabla u(t,\cdot) \in \mathcal{C}_b^\beta(\mathbb{R}^d)$ will be needed to construct a solution to (11) using Picard iterations, see Theorem 4.4, and $\nabla u(t,\cdot) \in \mathcal{C}_b^{\gamma_0/2}(\mathbb{R}^d)$ will be used when we apply Itô's formula, see Proposition A.2 and the proof of Theorem 2.1. Note that $\nabla u(t,\cdot) \in \mathcal{C}_b^{\gamma_0/2}(\mathbb{R}^d) \subseteq \mathcal{C}_b^{\gamma_0 - 1}(\mathbb{R}^d)$ implies, in particular, that $u(t,\cdot) \in \mathcal{D}(A)$ for all $t \in [0,T]$, see the remark following (7).

Proof of Lemma 4.2. We claim that

$$u(t,x) := \int_{(0,t)} \mathbb{E} g(s, x + L_{t-s})\,ds, \quad t \in [0,T],\ x \in \mathbb{R}^d,$$

has all the desired properties.

Step 1: u satisfies (14)**.** For fixed $i \in \{1,\dots,d\}$ and $t \in (0,T]$ define

$$F_t\phi(x) := \int_{\mathbb{R}^d} \phi(x+y)\partial_{y_i} p_t(y)\,dy = \int_{\mathbb{R}^d} \phi(y)\partial_{y_i} p_t(y-x)\,dy, \quad \phi \in C_b(\mathbb{R}^d),\ x \in \mathbb{R}^d.$$

By Lemma 4.1,

$$\frac{\partial}{\partial x_i} \mathbb{E}\phi(x + L_t) = \frac{\partial}{\partial x_i} \int \phi(y) p_t(y-x)\,dy = -F_t\phi(x)$$

and because of (3) we know that

$$\|F_t\phi\|_\infty \leq \|\phi\|_\infty \int_{\mathbb{R}^d} |\partial_{y_i} p_t(y)|\,dy \leq c\|\phi\|_\infty t^{-1/\alpha} \tag{15}$$

as well as

$$|F_t\phi(x) - F_t\phi(z)| = \left| \int_{\mathbb{R}^d} (\phi(x+y) - \phi(z+y))\partial_{y_i} p_t(y)\,dy \right| \leq c\|\phi\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} |x-z|^\beta t^{-1/\alpha}$$

for some absolute constant $c > 0$; therefore,

$$\|\partial_{x_i} \mathbb{E}\phi(\cdot + L_t)\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} = \|F_t\phi\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} \leq 2ct^{-1/\alpha}\|\phi\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} \quad \text{for all } \phi \in \mathcal{C}_b^\beta(\mathbb{R}^d).$$

Applying this to $\phi(y) := g(s,y)$ (with $s \in (0,t)$ fixed) it follows from the differentiation lemma and (3) that $\partial_{x_i} u(t,x)$ exists for all $t \in (0,T]$ and

$$\partial_{x_i} u(t,x) = \int_{(0,t)} \partial_{x_i} \mathbb{E} g(s, x + L_{t-s})\,ds$$

satisfies

$$\|\partial_{x_i} u\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} \leq 2c\|g\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} \int_0^T (T-s)^{-1/\alpha}\,ds =: C_1 T^{1-1/\alpha}\|g\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)};$$

(recall that, by assumption, $\alpha > 1$). In order to prove $\|\partial_{x_i} u\|_{\mathcal{C}_b^{\gamma_0/2}(\mathbb{R}^d)} \leq C_2(T)\|g\|_{\mathcal{C}_b^{\beta}(\mathbb{R}^d)}$ we show that

$$\|F_t\phi\|_{\mathcal{C}_b^{\gamma_0/2}(\mathbb{R}^d)} \leq c'\|\phi\|_{\mathcal{C}_b^{\beta}(\mathbb{R}^d)} t^{-(2+\gamma_0(1-\beta))/\alpha}, \quad \phi \in \mathcal{C}_b^{\beta}(\mathbb{R}^d), \tag{16}$$

for some absolute constant $c' > 0$. By the differentiation lemma, we have

$$\partial_{x_k} F_t\phi(x) = \int_{\mathbb{R}^d} \phi(y)\partial_{y_k}\partial_{y_i} p_t(y-x)\,dy$$

for all $\phi \in C_b(\mathbb{R}^d)$, and so, by (3) and Lemma 4.1(iii),

$$\|\partial_{x_k} F_t\phi\|_\infty \leq ct^{-2/\alpha}\|\phi\|_\infty \quad \text{for all} \quad k = 1, \ldots, d,\ \phi \in C_b(\mathbb{R}^d). \tag{17}$$

If $\gamma_0 \in [1,2)$, then we can use real interpolation to get $\mathcal{C}_b^{\gamma_0/2}(\mathbb{R}^d) = (C_b(\mathbb{R}^d), C_b^1(\mathbb{R}^d))_{\gamma_0/2,\infty}$, cf. Triebel [28, Section 2.7.2] or Lunardi [16, Example 1.8]. From (15), (17) and the interpolation theorem, see e.g. [28, Section 1.3.3] or [16, Theorem 1.6], it follows that

$$\|F_t\phi\|_{\mathcal{C}_b^{\gamma_0/2}(\mathbb{R}^d)} \leq \|F_t\phi\|_{C_b(\mathbb{R}^d)}^{1-\gamma_0/2}\|F_t\phi\|_{C_b^1(\mathbb{R}^d)}^{\gamma_0/2} \leq c\|\phi\|_\infty t^{-(1+\gamma_0/2)/\alpha}, \quad \phi \in C_b(\mathbb{R}^d). \tag{18}$$

If $\gamma_0 = 2$, then (18) is a direct consequence of (17). On the other hand, another application of the differentiation lemma shows that for any $\phi \in C_b^1(\mathbb{R}^d)$

$$\partial_{x_k} F_t\phi(x) = \int_{\mathbb{R}^d} \partial_{x_k}\phi(x+y)\partial_{y_i} p_t(y)\,dy$$

implying

$$\|\partial_{x_k} F_t\phi\|_\infty \leq c\|\phi\|_{C_b^1(\mathbb{R}^d)} t^{-1/\alpha}, \quad \phi \in C_b^1(\mathbb{R}^d).$$

Thus,

$$\|F_t\phi\|_{\mathcal{C}_b^{\gamma_0/2}(\mathbb{R}^d)} \leq c\|\phi\|_{C_b^1(\mathbb{R}^d)} t^{-1/\alpha}, \quad \phi \in C_b^1(\mathbb{R}^d). \tag{19}$$

Using (18) and (19) we can apply the interpolation theorem once more to find that F_t maps $\mathcal{C}_b^{\beta}(\mathbb{R}^d) = (C_b(\mathbb{R}^d), C_b^1(\mathbb{R}^d))_{\beta,\infty}$ into $\mathcal{C}_b^{\gamma_0/2}(\mathbb{R}^d)$ and

$$\|F_t\phi\|_{\mathcal{C}_b^{\gamma_0/2}(\mathbb{R}^d)} \leq c\|\phi\|_{\mathcal{C}_b^{\beta}(\mathbb{R}^d)} t^{-(1-\beta)(1+\gamma_0/2)/\alpha - \beta/\alpha} = c\|\phi\|_{\mathcal{C}_b^{\beta}(\mathbb{R}^d)} t^{-\kappa}.$$

for $\kappa := (2+\gamma_0(1-\beta))/(2\alpha)$; note that $\kappa < 1$ because of the balance condition (4). Applying the estimate to $\phi(y) := g(s,y)$, we conclude that

$$\|\partial_{x_i} u\|_{\mathcal{C}_b^{\gamma_0/2}(\mathbb{R}^d)} \leq c\|g\|_{\mathcal{C}_b^{\beta}(\mathbb{R}^d)} \int_0^T (T-s)^{-\kappa}\,ds =: C_2 T^{1-\kappa}\|g\|_{\mathcal{C}_b^{\beta}(\mathbb{R}^d)}.$$

Step 2: u solves (13). By [13, Theorem 4.1(iii)], we have $\mathcal{C}_\infty^{1,\gamma_0-1}(\mathbb{R}^d) \subseteq \mathcal{D}(A)$. It follows from the proof of Step 1 that

$$x \mapsto P_\epsilon\phi(x) := \mathbb{E}\phi(x+L_\epsilon) \in \mathcal{C}_\infty^{1,\gamma_0/2}(\mathbb{R}^d) \subseteq \mathcal{C}_\infty^{1,\gamma_0-1}(\mathbb{R}^d) \subseteq \mathcal{D}(A)$$

for all $\epsilon > 0$ and $\phi \in \mathcal{C}_b^{\beta}(\mathbb{R}^d) \cap C_\infty(\mathbb{R}^d)$. Since

$$\frac{d}{dt}P_t f = AP_t f \quad \text{for all} \quad f \in \mathcal{D}(A)$$

we find

$$\frac{d}{dt}P_{t+\epsilon}\phi = AP_{t+\epsilon}\phi$$

for all $t \geq 0$, $\epsilon > 0$ and $\phi \in \mathcal{C}_b^{\beta}(\mathbb{R}^d) \cap C_\infty(\mathbb{R}^d)$ which means that

$$\frac{d}{d\tau}P_\tau\phi = AP_\tau\phi, \quad \tau > 0,\ \phi \in \mathcal{C}_b^{\beta}(\mathbb{R}^d) \cap C_\infty(\mathbb{R}^d).$$

Applying this identity to $\phi(x) := g(s,x)$ with $g \in C([0,T], \mathcal{C}_b^{\beta}(\mathbb{R}^d))$ such that $g(s,\cdot) \in C_\infty(\mathbb{R}^d)$ for all $s \in [0,T]$ shows that $u(t,x) = \int_{(0,t)} \mathbb{E}g(s, x+L_{t-s})\,ds$ is a function in $C^1([0,T], C_b(\mathbb{R}^d))$ which solves (13), see [3, Lemma 7] for details.

For an arbitrary $g \in C([0,T], \mathcal{C}_b^\beta(\mathbb{R}^d))$ fix a cut-off function $\chi \in C_c^\infty(\mathbb{R}^d)$ such that $\mathbb{1}_{B(0,1)} \leq \chi \leq \mathbb{1}_{B(0,2)}$ and set $g_n(t,x) := g(t,x)\chi(x/n)$ for $t \in [0,T]$, $x \in \mathbb{R}^d$ and $n \in \mathbb{N}$. Since $g_n(t,\cdot)$ vanishes at infinity, it follows from the first part that $u_n(t,x) := \int_{(0,t)} \mathbb{E} g_n(s, x + L_{t-s})\, ds$ satisfies

$$\frac{\partial}{\partial t} u_n(t,x) = A_x u_n(t,x) + g_n(t,x), \quad u_n(0,x) = 0$$

i.e.

$$u_n(t,x) = \int_{(0,t)} (A_x u_n(s,x) + g_n(s,x))\, ds. \tag{20}$$

We are going to show that we can let $n \to \infty$ using the dominated convergence theorem. For any $R > 0$ and $t \in [0,T]$ we have

$$|u_n(t,x) - u(t,x)| \leq \sup_{s \in [0,T]} \sup_{|y+x| \leq R} |g_n(s,y) - g(s,y)| + 2T\|g\|_\infty \mathbb{P}\left(\sup_{0 \leq s \leq T} |L_s + x| > R \right)$$

and, by Step 1,

$$\begin{aligned} |\partial_{x_i} u_n(t,x) - \partial_{x_i} u(t,x)| &\leq \int_0^t \int_{|y| \leq R} |g_n(s,y) - g(s,y)| \cdot |\partial_{y_i} p_{t-s}(y-x)|\, dy\, ds \\ &\quad + 2\|g\|_\infty \int_0^t \int_{|y|>R} |\partial_{y_i} p_{t-s}(y-x)|\, dy\, ds. \end{aligned}$$

Therefore, we can combine the dominated convergence theorem, (3) and the fact that $g_n \to g$ converges uniformly on compact sets to see that

$$\sup_{t \in [0,T]} |u(t,x) - u_n(t,x)| + \sup_{t \in [0,T]} |\nabla_x u(t,x) - \nabla_x u_n(t,x)| \xrightarrow{n \to \infty} 0 \tag{21}$$

for all $x \in \mathbb{R}^d$. Moreover, by Step 1,

$$\begin{aligned} \|u_n\|_{\mathcal{C}_b^{1,\gamma_0-1}(\mathbb{R}^d)} + \|u\|_{\mathcal{C}_b^{1,\gamma_0-1}(\mathbb{R}^d)} &\leq \|u_n\|_{\mathcal{C}_b^{1,\gamma_0/2}(\mathbb{R}^d)} + \|u\|_{\mathcal{C}_b^{1,\gamma_0/2}(\mathbb{R}^d)} \\ &\leq C\|g\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} + C\|g_n\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} \leq 2C\|g\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)}. \end{aligned}$$

Observe that we have $\int_{|y| \leq 1} |y|^{\gamma_0}\, \nu(dy) < \infty$ and for any function $f \in \mathcal{C}_b^{1,\gamma_0-1}(\mathbb{R}^d)$

$$|f(x+y) - f(x)| \leq 2\|f\|_\infty \quad \text{and} \quad |f(x+y) - f(x) - \nabla f(x) \cdot y| \leq \|f\|_{\mathcal{C}_b^{1,\gamma_0-1}(\mathbb{R}^d)} |y|^{\gamma_0}.$$

Therefore, we can use dominated convergence and (21) to infer that

$$\sup_{t \in [0,T]} |A_x u_n(t,x) + A_x u(t,x)| \xrightarrow{n \to \infty} 0 \quad \text{for all} \quad x \in \mathbb{R}^d.$$

Letting $n \to \infty$ in (20), we finally get

$$u(t,x) = \int_{(0,t)} (A_x u(s,x) + g(s,x))\, ds, \quad t \in [0,T],\ x \in \mathbb{R}^d.$$

This shows that $u \in C^1([0,T], C_b(\mathbb{R}^d))$ solves (13). □

4.4. **Theorem.** *Let $(L_t)_{t \geq 0}$ be a d-dimensional Lévy process as in Theorem 2.1 with infinitesimal generator $(A, \mathcal{D}(A))$, and $b, g \in C([0,T], \mathcal{C}_b^\beta(\mathbb{R}^d))$ for $\beta \in (0,1]$ satisfying (4). For sufficiently small $T > 0$ there exists a map $u \in C([0,T], \mathcal{C}_b^{1,\max\{\beta,\gamma_0/2\}}(\mathbb{R}^d)) \cap C^1([0,T], C_b(\mathbb{R}^d))$ solving the equation*

$$\begin{aligned} \frac{\partial}{\partial t} u(t,x) + A_x u(t,x) + b(t,x) \cdot \nabla_x u(t,x) &= -g(t,x) \quad \text{on } [0,T) \times \mathbb{R}^d \\ u(T,\cdot) &= 0. \end{aligned} \tag{22}$$

Moreover, u satisfies

$$\|u\|_\infty + \|\nabla_x u\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} + \|\nabla_x u\|_{\mathcal{C}_b^{\gamma_0/2}(\mathbb{R}^d)} \leq c(T)\|g\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} \tag{23}$$

for some constant $c(T) > 0$ which does not depend on b, g and $c(T) \to 0$ as $T \to 0$.

Proof. We use Picard iteration to prove the existence of the solution. Choose $T > 0$ so small that $2C(T)\|b\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} \le 1/2$ where $C(T)$ is the constant appearing in (14), and set $u^{(0)} := 0$. By Lemma 4.2 and Remark 4.3(i) we can define iteratively $u^{(n+1)} \in C([0,T], \mathcal{C}_b^{\max\{\beta,\, \gamma_0/2\}}(\mathbb{R}^d)) \cap C^1([0,T], C_b(\mathbb{R}^d))$ such that (14) holds and

$$\frac{\partial}{\partial t}u^{(n+1)}(t,x) + A_x u^{(n+1)}(t,x) = -b(t,x)\cdot\nabla_x u^{(n)}(t,x) - g(t,x), \quad u^{(n+1)}(T,\cdot) = 0.$$

Using repeatedly (14) we find

$$\begin{aligned}\|u^{(n+1)} - u^{(n)}\|_{\mathcal{C}_b^{1,\max\{\beta,\,\gamma_0/2\}}(\mathbb{R}^d)} &\le C(T)\|b\cdot\nabla_x u^{(n)} - b\cdot\nabla_x u^{(n-1)}\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)}\\ &\le 2C(T)\|b\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)}\|\nabla_x u^{(n)} - \nabla_x u^{(n-1)}\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)}\\ &\le \frac{1}{2}\|\nabla_x u^{(n)} - \nabla_x u^{(n-1)}\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} \le \cdots \le \frac{1}{2^n}\|g\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)},\end{aligned}$$

and, therefore,

$$\sum_{n\ge1}\|u^{(n+1)} - u^{(n)}\|_{\mathcal{C}_b^{1,\max\{\beta,\,\gamma_0/2\}}(\mathbb{R}^d)} < \infty.$$

Since $C([0,T], \mathcal{C}_b^{1,\max\{\beta,\,\gamma_0/2\}}(\mathbb{R}^d))$ is a Banach space, completeness implies that there is some $u \in C([0,T], \mathcal{C}_b^{1,\max\{\beta,\,\gamma_0/2\}}(\mathbb{R}^d))$ such that $u^{(n)} \to u$ in $C([0,T], \mathcal{C}_b^{1,\max\{\beta,\,\gamma_0/2\}}(\mathbb{R}^d))$. In particular, by (8),

$$\|Au(t,\cdot) - Au^{(n)}(t,\cdot)\|_\infty \le M\|u - u^{(n)}\|_{\mathcal{C}_b^{1,\gamma_0-1}(\mathbb{R}^d)} \le M\|u - u^{(n)}\|_{\mathcal{C}_b^{1,\gamma_0/2}(\mathbb{R}^d)} \xrightarrow{n\to\infty} 0$$

(note that $\gamma_0 - 1 \le \gamma_0/2$ as $\gamma_0 \in [1,2]$). Letting $n \to \infty$ in

$$u^{(n)}(t,x) = \int_t^T \Big(A_x u^{(n)}(s,x) + b(s,x)\cdot\nabla_x u^{(n)}(s,x) + g(s,x)\Big)\, ds$$

we get

$$u(t,x) = \int_t^T (A_x u(s,x) + b(s,x)\cdot\nabla_x u(s,x) + g(s,x))\, ds.$$

Using the above estimates, it is not difficult to see that u has all the desired properties. □

We are now ready to prove Theorem 2.1.

Proof of Theorem 2.1. By considering each coordinate of $X_t \in \mathbb{R}^d$ separately, we may assume, without loss of generality, that $d = 1$. Fix some sufficiently small $\epsilon > 0$ (we will specify ϵ later in the proof), $p \le \gamma_\infty$, $T > 0$ and set $T_i := T\frac{i}{L}$, $i = 0, \dots, L$.

If we choose $L = L(\epsilon) \in \mathbb{N}$ sufficiently large, Theorem 4.4 shows that there exists a function $u_i \in C([T_{i-1}, T_i], \mathcal{C}_b^{1,\max\{\beta,\,\gamma_0/2\}}(\mathbb{R})) \cap C^1([T_{i-1}, T_i], C_b(\mathbb{R}))$ such that

$$\begin{aligned}\frac{\partial}{\partial t}u_i(t,x) + A_x u_i(t,x) + b(t,x)\frac{\partial}{\partial x}u_i(t,x) &= -b(t,x) \quad \text{on } [T_{i-1}, T_i)\times\mathbb{R}\\ u_i(T_i,\cdot) &= 0\end{aligned} \tag{24}$$

and

$$\|u_i\|_\infty + \|\partial_x u_i\|_{\mathcal{C}_b^\beta(\mathbb{R}^d)} + \|\partial_x u_i\|_{\mathcal{C}_b^{\gamma_0/2}(\mathbb{R}^d)} \le \epsilon. \tag{25}$$

Denote by $(X_t^{(n)})_{t\ge0}$ the Euler–Maruyama approximation, i.e.

$$X_t^{(n)} = x + \int_0^t b(\eta_n(s), X^{(n)}_{\eta_n(s)-})\, ds + L_t$$

where $\eta_n(s) := T\frac{i}{n}$ for $s \in [T\frac{i}{n}, T\frac{i+1}{n})$. We are going to show that

$$\mathbb{E}\left(\sup_{0\le t\le T}|X_t^{(n)} - X_t^{(m)}|^p\right) \xrightarrow{m,n\to\infty} 0.$$

Applying Itô's formula, cf. Proposition A.2, it follows from (24) that

$$
\begin{aligned}
\int_{T_{i-1}}^t b(s, X_s^{(n)})\, ds &= u_i(T_{i-1}, X_{T_{i-1}}^{(n)}) - u_i(t, X_t^{(n)}) \\
&\quad - \int_{T_{i-1}}^t \int_{|y|<1} \big(u_i(s, X_{s-}^{(n)} + y) - u_i(s, X_{s-}^{(n)})\big)\, \tilde{N}(dy, ds) \\
&\quad - \int_{T_{i-1}}^t \int_{|y|\geq 1} \big(u_i(s, X_{s-}^{(n)} + y) - u_i(s, X_{s-}^{(n)})\big)\, N(dy, ds) \\
&\quad - \int_{T_{i-1}}^t \big(b(\eta_n(s), X_{\eta_n(s)}^{(n)}) - b(s, X_s^{(n)})\big)\partial_x u_i(s, X_s^{(n)})\, ds
\end{aligned} \tag{26}
$$

for any $t \in [T_{i-1}, T_i)$, $i = 0, \dots, L$, where $\tilde{N}(dy, ds) = N(dy, ds) - \nu(dy)\, ds$ denotes the compensated jump measure of the Lévy process $(L_t)_{t\geq 0}$. Fix $i \in \{0, \dots, L\}$, $t \in [T_{i-1}, T_i]$ and $m, n \in \mathbb{N}$. Observing that

$$
|X_t^{(m)} - X_t^{(n)}| \leq |X_{T_{i-1}}^{(m)} - X_{T_{i-1}}^{(n)}| + \left| \int_{T_{i-1}}^t b(\eta_m(s), X_{\eta_m(s)}^{(m)})\, ds - \int_{T_{i-1}}^t b(\eta_n(s), X_{\eta_n(s)}^{(n)})\, ds \right|,
$$

we get

$$
|X_t^{(m)} - X_t^{(n)}|^p \leq C|X_{T_{i-1}}^{(m)} - X_{T_{i-1}}^{(n)}|^p + C(I_1 + I_2 + I_3 + I_{4,1} + I_{4,2} + I_5)
$$

for some constant $C = C(p) > 0$ and the following (integral) expressions

$$
\begin{aligned}
I_1 &:= \left| u_i(T_{i-1}, X_{T_{i-1}}^{(m)}) - u_i(T_{i-1}, X_{T_{i-1}}^{(n)}) \right|^p + \left| u_i(t, X_t^{(m)}) - u_i(t, X_t^{(n)}) \right|^p \\
I_2 &:= \left| \int_{T_{i-1}}^t \int_{|y|<1} H_i(s, y)\, \tilde{N}(dy, ds) \right|^p \\
I_3 &:= \left| \int_{T_{i-1}}^t \int_{|y|\geq 1} H_i(s, y)\, N(dy, ds) \right|^p \\
I_{4,1} &:= \left| \int_{T_{i-1}}^t \big(b(\eta_n(s), X_{\eta_n(s)}^{(n)}) - b(s, X_s^{(n)})\big)\partial_x u_i(s, X_s^{(n)})\, ds \right|^p \\
I_{4,2} &:= \left| \int_{T_{i-1}}^t \big(b(\eta_m(s), X_{\eta_m(s)}^{(m)}) - b(s, X_s^{(m)})\big)\partial_x u_i(s, X_s^{(m)})\, ds \right|^p \\
I_5 &:= \left| \int_{T_{i-1}}^t (b(\eta_n(s), X_{\eta_n(s)}^{(n)}) - b(s, X_s^{(n)})\, ds \right|^p + \left| \int_{T_{i-1}}^t (b(\eta_m(s), X_{\eta_m(s)}^{(m)}) - b(s, X_s^{(m)})\, ds \right|^p
\end{aligned}
$$

and

$$
H_i(s, y) := \big(u_i(s, X_{s-}^{(m)} + y) - u_i(s, X_{s-}^{(m)})\big) - \big(u_i(s, X_{s-}^{(n)} + y) - u_i(s, X_{s-}^{(n)})\big).
$$

We estimate the terms separately. Because of (25), we have $\|\partial_x u_i\|_\infty \leq \epsilon$, and therefore an application of the mean value theorem shows

$$
I_1 \leq \epsilon^p \left| X_{T_{i-1}}^{(m)} - X_{T_{i-1}}^{(n)} \right|^p + \epsilon^p \left| X_t^{(m)} - X_t^{(n)} \right|^p.
$$

Moreover, it follows from (25) and the fact that $b(t, x)$ is β-Hölder-continuous with respect to x and η-Hölder continuous with respect to t that

$$
\begin{aligned}
I_{4,1} &\leq C \left| \int_{T_{i-1}}^t \big(b(\eta_n(s), X_{\eta_n(s)}^{(n)}) - b(s, X_{\eta_n(s)}^{(n)})\big)\partial_x u_i(s, X_s^{(n)})\, ds \right|^p \\
&\quad + C \left| \int_{T_{i-1}}^t \big(b(s, X_{\eta_n(s)}^{(n)}) - b(s, X_s^{(n)})\big)\partial_x u_i(s, X_s^{(n)})\, ds \right|^p \\
&\leq C_4 n^{-p\eta} \epsilon^p + C_4' \epsilon^p \sup_{T_{i-1} \leq t \leq T_i} |X_{\eta_n(t)}^{(n)} - X_t^{(n)}|^{\beta p}.
\end{aligned}
$$

The same estimate holds for $I_{4,2}$ with n replaced by m. In exactly the same fashion we get

$$I_5 \le C_5 n^{-\eta p} + C_5' \epsilon^p \sup_{T_{i-1}\le t\le T_i} |X^{(n)}_{\eta_n(t)} - X^{(n)}_t|^{\beta p} + C_5' \epsilon^p \sup_{T_{i-1}\le t\le T_i} |X^{(m)}_{\eta_m(t)} - X^{(m)}_t|^{\beta p}.$$

In order to estimate I_2 and I_3, we use Taylor's formula and (25)

$$\begin{aligned}|H_i(s,y)| &\le |X^{(m)}_{s-} - X^{(n)}_{s-}| \int_0^1 \bigg| \frac{\partial}{\partial x} u_i\big(s, X^{(m)}_s + y + \tau(X^{(n)}_{s-} - X^{(m)}_{s-})\big) \\ &\qquad\qquad - \frac{\partial}{\partial x} u_i\big(s, X^{(m)}_{s-} + \tau(X^{(n)}_{s-} - X^{(m)}_{s-})\big)\bigg|\, d\tau \\ &\le \epsilon \min\big\{ |X^{(m)}_{s-} - X^{(n)}_{s-}|\, |y|^{\gamma_0/2},\ 2|X^{(m)}_{s-} - X^{(n)}_{s-}| \big\}.\end{aligned} \tag{27}$$

Applying the Burkholder–Davis–Gundy inequality, cf. Novikov [19, Theorem 1], we find

$$\begin{aligned}&\mathbb{E}\left(\sup_{T_{i-1}\le t\le T_i} |I_2| \right) \\ &\le C_2 \mathbb{E}\left[\left(\int_{T_{i-1}}^{T_i} \int_{|y|<1} |H_i(s,y)|^2\, \nu(dy)\, ds \right)^{p/2} \right] + C_2 \mathbb{E}\left(\int_{T_{i-1}}^{T_i} \int_{|y|<1} |H_i(s,y)|^p\, \nu(dy)\, ds \right) \mathbb{1}_{[2,\infty)}(p) \\ &\le C_2' \epsilon^p \left(\left[\int_{|y|<1} |y|^{\gamma_0}\, \nu(dy) \right]^{p/2} + \int_{|y|<1} |y|^{\gamma_0}\, \nu(dy) \right) \mathbb{E}\left(\sup_{T_{i-1}\le t\le T_i} |X^{(m)}_s - X^{(n)}_s|^p \right)\end{aligned}$$

for some absolute constants $C_2, C_2' > 0$. In order to estimate I_3 we distinguish between two cases. If $p \in (0,1)$, then $(x+y)^p \le x^p + y^p$ for all $x, y \ge 0$, and therefore by (27)

$$\begin{aligned}\mathbb{E}\left(\sup_{T_{i-1}\le t\le T_i} |I_3| \right) &= \mathbb{E}\left(\sup_{T_{i-1}\le t\le T_i} \bigg| \sum_{\substack{s\in[T_{i-1},t]\\ |\Delta L_s|\ge 1}} H_i(s, \Delta L_s) \bigg|^p \right) \\ &\le \mathbb{E}\left(\sum_{\substack{s\in[T_{i-1},T_i]\\ |\Delta L_s|\ge 1}} |H_i(s, \Delta L_s)|^p \right) \\ &= \mathbb{E}\left(\int_{T_{i-1}}^{T_i} \int_{|y|\ge 1} |H_i(s,y)|^p\, N(dy, ds) \right) \\ &= \mathbb{E}\left(\int_{T_{i-1}}^{T_i} \int_{|y|\ge 1} |H_i(s,y)|^p\, \nu(dy)\, ds \right) \\ &\le 2^p \epsilon^p \nu(B(0,1)^c) \mathbb{E}\left(\sup_{T_{i-1}\le t\le T_i} |X^{(m)}_t - X^{(n)}_t|^p \right).\end{aligned}$$

If $p \ge 1$, then $\int_{|y|\ge 1} |y|\, \nu(dy) < \infty$, and so

$$I_3 \le C \left| \int_{T_{i-1}}^t \int_{|y|\ge 1} H_i(s,y)\, \tilde N(dy, ds) \right|^p + C \left| \int_{T_{i-1}}^t \int_{|y|\ge 1} H_i(s,y)\, \nu(dy)\, ds \right|^p.$$

By the Burkholder–Davis–Gundy inequality and (27), there exist absolute constants $C_3, C_3' > 0$ such that

$$\begin{aligned}&\mathbb{E}\left(\sup_{T_{i-1}\le t\le T_i} I_3 \right) \\ &\le C_3 \mathbb{E}\left(\left[\int_{T_{i-1}}^{T_i} \int_{|y|\ge 1} |H_i(s,y)|^2\, \nu(dy)\, ds \right]^{p/2} \right) + C_3 \mathbb{E}\left(\int_{T_{i-1}}^{T_i} \int_{|y|\ge 1} |H_i(s,y)|^p\, \nu(dy)\, ds \right) \\ &\le C_3' \epsilon^p \big(\nu(B(0,1)^c)^{p/2} + \nu(B(0,1)^c) \big) \mathbb{E}\left(\sup_{T_{i-1}\le t\le T_i} |X^{(m)}_t - X^{(n)}_t|^p \right).\end{aligned}$$

Combining the above estimates we conclude that there exist constants $c_1, c_2 > 0$ (not depending on ϵ, m, n, L, i) such that

$$\begin{aligned}
&\mathbb{E}\left(\sup_{T_{i-1}\le t\le T_i} |X_t^{(m)} - X_t^{(n)}|^p\right) \\
&\quad\le \epsilon^p c_1 \mathbb{E}\left(\sup_{T_{i-1}\le t\le T_i} |X_t^{(m)} - X_t^{(n)}|^p\right) + c_2\mathbb{E}\left(|X_{T_{i-1}}^{(m)} - X_{T_{i-1}}^{(n)}|^p\right) + \frac{c_2}{n^{p\eta}} + \frac{c_2}{m^{p\eta}} \\
&\qquad + c_2\mathbb{E}\left(\sup_{T_{i-1}\le t\le T_i} |X_{\eta_n(t)}^{(n)} - X_t^{(n)}|^{p\beta}\right) + c_2\mathbb{E}\left(\sup_{T_{i-1}\le t\le T_i} |X_{\eta_m(t)}^{(m)} - X_t^{(m)}|^{p\beta}\right).
\end{aligned}$$

Thus,

$$\begin{aligned}
&(1-\epsilon^p c_1)\mathbb{E}\left(\sup_{T_{i-1}\le t\le T_i} |X_t^{(m)} - X_t^{(n)}|^p\right) \\
&\quad\le c_2\mathbb{E}\left(|X_{T_{i-1}}^{(m)} - X_{T_{i-1}}^{(n)}|^p\right) + \frac{2c_2}{N^{p\eta}} + 2c_2 \sup_{n\ge N} \mathbb{E}\left(\sup_{T_{i-1}\le t\le T_i} |X_{\eta_n(t)}^{(n)} - X_t^{(n)}|^{p\beta}\right)
\end{aligned}$$

for any $m, n \ge N$. Choose $\epsilon > 0$ so small that $1 - \epsilon^p c_1 \ge 1/2$. By the very definition of the Euler–Maruyama approximation, we have

$$X_{\eta_n(t)}^{(n)} - X_t^{(n)} = \int_{\eta_n(t)}^t b(\eta_n(s), X_{\eta_n(s)-}^{(n)})\, ds + L_{\eta_n(t)} - L_t.$$

Using $L_t - L_{\eta_n(t)} \overset{d}{=} L_{t-\eta_n(t)}$ and fractional moment estimates for Lévy processes, see [10, Section 5], we can find a constant $c_3 > 0$ such that

$$\begin{aligned}
\mathbb{E}\left(\sup_{T_{i-1}\le t\le T_i} |X_{\eta_n(t)}^{(n)} - X_t^{(n)}|^{p\beta}\right) &\le (2\|b\|_\infty)^{p\beta} n^{-p\beta} + 2^{p\beta}\mathbb{E}\left(\sup_{T_{i-1}\le t\le T_i} |L_{t-\eta_n(t)}|^{p\beta}\right) \\
&\le (2\|b\|_\infty)^{p\beta} n^{-p\beta} + 2^{p\beta}\mathbb{E}\left(\sup_{s\le 1/n} |L_s|^{p\beta}\right) \\
&\le c_3 n^{-\min\{1,\, p\beta/\gamma_0\}}.
\end{aligned}$$

Hence,

$$\mathbb{E}\left(\sup_{T_{i-1}\le t\le T_i} |X_t^{(m)} - X_t^{(n)}|^p\right) \le 2c_2\mathbb{E}\left(|X_{T_{i-1}}^{(m)} - X_{T_{i-1}}^{(n)}|^p\right) + 8c_2c_3 N^{-\min\{1, p\beta/\gamma_0, p\eta\}}.$$

Using this estimate iteratively for $i = 1, \dots, L$, we conclude that there exists a constant $c_4 = c_4(L) > 0$ such that

$$\mathbb{E}\left(\sup_{T_{i-1}\le t\le T_i} |X_t^{(m)} - X_t^{(n)}|^p\right) \le c_4 N^{-\min\{1, p\beta/\gamma_0, p\eta\}}$$

for all $m, n \ge N$. Thus,

$$\mathbb{E}\left(\sup_{0\le t\le T} |X_t^{(m)} - X_t^{(n)}|^p\right) \le \sum_{i=1}^L \mathbb{E}\left(\sup_{T_{i-1}\le t\le T_i} |X_t^{(m)} - X_t^{(n)}|^p\right) \le c_4 L N^{-\min\{1, p\beta/\gamma_0, p\eta\}} \tag{28}$$

for all $m, n \ge N$; this means, in particular, that

$$\mathbb{E}\left(\sup_{0\le t\le T} |X_t^{(m)} - X_t^{(n)}|^p\right) \xrightarrow{m,n\to\infty} 0.$$

This implies that there exists a stochastic process $(X_t)_{t\in[0,T]}$ and a subsequence $(n_k)_{k\in\mathbb{N}}$ such that $\sup_{t\in[0,T]} |X_t^{(n_k)} - X_t| \to 0$ almost surely as $k \to \infty$, cf. Lemma A.3. Letting $k \to \infty$ in

$$X_t^{(n_k)} - x = \int_0^t b(\eta_{n_k}(s), X_{\eta_{n_k}(s)}^{(n_k)})\, ds + L_t$$

we find

$$X_t - x = \int_0^t b(s, X_{s-})\,ds + L_t.$$

Moreover, it follows from Fatou's lemma and (28) that

$$\mathbb{E}\left(\sup_{0\le t\le T} |X_t - X_t^{(n)}|^p\right) \le c_4' n^{-\min\{1,\,p\beta/\gamma_0,\,p\eta\}}.$$

This proves the existence of a solution to (5) satisfying (6).

In order to show uniqueness, we assume that $(Y_t)_{t\ge 0}$ is a further solution to (5). Applying Itô's formula to $u_i(t, Y_t)$ with u_i as in (24), we get a similar expression as in (26) for $\int_0^t b(s, Y_s)\,ds$, and a very similar reasoning as in the first part of the proof to shows that

$$\mathbb{E}\left(\sup_{0\le t\le T} |Y_t - X_t^{(n)}|^p\right) \le c_5 n^{-\min\{1,\,p\beta/\gamma_0,\,p\eta\}} \tag{29}$$

for all $n \in \mathbb{N}$. Thus, by Fatou's lemma,

$$\mathbb{E}\left(\sup_{0\le t\le T} |Y_t - X_t|^p\right) = 0. \qquad \square$$

Proof of Corollary 2.4. From [26, Theorem 1.3] it follows that the semigroup $P_t\phi(x) := \mathbb{E}\phi(x + L_t)$ satisfies

$$\|\nabla P_t \phi\|_\infty \le c t^{-1/\alpha} \|\phi\|_\infty \quad \text{for all} \quad \phi \in \mathcal{B}_b(\mathbb{R}^d).$$

By Lemma 4.1, this implies $\int_{\mathbb{R}^d} |\partial_i p_t(x)|\,dx \le c t^{-1/\alpha}$ for all $i \in \{1, \dots, d\}$. Applying Theorem 2.1 finishes the proof. $\square$

The remaining part of this section is devoted to the proof of Corollary 2.5. From now on $(S_t)_{t\ge 0}$ denotes a subordinator with Laplace exponent (Bernstein function) f, $(B_t^{(d)})_{t\ge 0}$ is a d-dimensional Brownian motion and $L_t^{(d)} := B_{S_t}^{(d)}$ is the process subordinate to Brownian motion. Note that $(L_t^{(d)})_{t\ge 0}$ is a d-dimensional Lévy process with characteristic exponent $\psi(\xi) = f(|\xi|^2)$, $\xi \in \mathbb{R}^d$. If f satisfies the Hartman–Wintner condition

$$\lim_{r\to\infty} \frac{f(r)}{\log(1+r)} = \infty, \tag{30}$$

then $L_t^{(d)}$ has for all $t > 0$ a transition density $p_t^{(d)}$ see e.g. [8], and $p_t^{(d)}$ is isotropic, i.e. $p_t^{(d)}(x)$ depends only on $|x|$; in abuse of notation we write $p_t^{(d)}(x) = p_t^{(d)}(|x|)$. Using polar coordinates one finds, cf. Matheron [17, pp. 33–4] and [14],

$$\frac{d}{dr} p_t^{(d)}(r) = -2\pi r p_t^{(d+2)}(r), \quad r > 0,\ d \ge 1. \tag{31}$$

4.5. **Lemma.** *Let f be a Bernstein function satisfying the Hartman–Wintner condition* (30)*, and let $(L_t^{(d)})_{t\ge 0}$ be a d-dimensional Lévy process with characteristic exponent $\psi(\xi) = f(|\xi|^2)$, $\xi \in \mathbb{R}^d$, for $d \ge 1$. If there exist constants $c > 0$, $\alpha > 0$ such that*

$$\mathbb{E}\left(\left|L_t^{(d+2)}\right|^{-1}\right) \le c t^{-1/\alpha} \quad \textit{for all} \quad t \in (0, T], \tag{32}$$

then the transition density of $L_t^{(d)}$, $t > 0$, satisfies (3).

Proof. Denote by $p_t^{(d)}(x) = p_t^{(d)}(|x|)$, $x \in \mathbb{R}^d$, the transition density of $(L_t^{(d)})_{t\geq 0}$. Using polar coordinates and (31), we find for each $i = 1, \dots, d$,

$$\begin{aligned}\int_{\mathbb{R}^d} |\partial_{x_i} p_t^{(d)}(x)|\,dx \leq 2\pi \int_{\mathbb{R}^d} |x| p_t^{(d+2)}(|x|)\,dx &= 2\pi\sigma_d \int_{(0,\infty)} r p_t^{(d+2)}(r) r^{d-1}\,dr \\ &= 2\pi\sigma_d \int_{(0,\infty)} \frac{1}{r} p_t^{(d+2)}(r) r^{(d+2)-1}\,dr \\ &= \frac{\sigma_d}{\sigma_{d+2}} \int_{\mathbb{R}^{d+2}} \frac{1}{|x|} p_t^{(d+2)}(|x|)\,dx \\ &= \frac{\sigma_d}{\sigma_{d+2}} \mathbb{E}\left(\left|L_t^{(d+2)}\right|^{-1}\right)\end{aligned}$$

where σ_d is the surface volume of the unit sphere $\mathbb{S}^{d-1} \subseteq \mathbb{R}^d$. Hence, by (32),

$$\int_{\mathbb{R}^d} |\partial_{x_i} p_t(x)|\,dx \leq c' t^{-1/\alpha} \quad \text{for all} \quad t \in (0,T],\ i = 1, \dots, d. \qquad \square$$

4.6. **Remark.** More generally, the condition

$$\mathbb{E}\left(\left|L_t^{(d+2i)}\right|^{-i}\right) \leq c t^{-i/\alpha}, \quad i = 1, \dots, k,\ t \in (0,T]$$

guarantees that

$$\int_{\mathbb{R}^d} |\partial_x^\gamma p_t(x)|\,dx \leq c' t^{-|\gamma|/\alpha} \quad \text{for all} \quad \gamma \in \mathbb{N}_0^d,\ |\gamma| \leq k,\ t \in (0,T].$$

Proof of Corollary 2.5. By assumption, there exists some $c > 0$ such that $\psi(\xi) \geq c|\xi|^{2\rho}$ for large $|\xi|$, and therefore

$$p_t(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{-ix\cdot\xi} e^{-t\psi(\xi)}\,d\xi, \quad x \in \mathbb{R}^d$$

is the density of L_t; by the differentiation lemma, p_t is twice continuously differentiable. To prove that the density of $L_t = L_t^{(d)} = B_{S_t}^{(d)}$, $t > 0$, satisfies (3) for $\alpha := 2\rho$, it suffices by Lemma 4.5 to show that (32) holds for $\alpha = 2\rho$. To this end, we recall that for any $\kappa > 0$ there exists a constant $C > 0$ such that

$$\mathbb{E}(S_t^{-\kappa}) \leq C \min\{t, 1\}^{-\kappa/\rho}, \quad t \geq 0, \tag{33}$$

cf. [2, Theorem 3.17]. As $(B_t^{(d)})_{t\geq 0}$ and $(S_t)_{t\geq 0}$ are independent, we get by the scaling property of Brownian motion

$$\mathbb{E}\left(\left|L_t^{(d+2)}\right|^{-1}\right) = \mathbb{E}\left(\left|B_{S_t}^{(d+2)}\right|^{-1}\right) = \mathbb{E}\left(\left|\sqrt{S_t} B_1^{(d+2)}\right|^{-1}\right) = \mathbb{E}\left(S_t^{-1/2}\right) \mathbb{E}\left(|B_1^{(d+2)}|^{-1}\right).$$

Note that

$$\mathbb{E}\left(\left|B_1^{(d+2)}\right|^{-1}\right) = \int_{\mathbb{R}^{d+2}} \frac{1}{|z|} \frac{1}{(2\pi)^{(d+2)/2}} \exp\left(-\frac{|z|^2}{2}\right) dz < \infty$$

as $1 < d + 2$. Because of (33) we get (32), hence (3), for $\alpha = 2\rho$.

Finally, since

$$\int_{(0,1)} r^{\delta_0}\,\mu(dr) + \int_{(1,\infty)} r^{\delta_\infty}\,\mu(dr) < \infty$$

implies

$$\int_{B(0,1)} |y|^{2\delta_0}\,\nu(dy) + \int_{B(0,1)^c} |y|^{2\delta_\infty}\,\nu(dy) < \infty,$$

the assumptions of Theorem 2.1 are satisfied for $\alpha := 2\rho$, $\gamma_0 := 2\delta_0$, $\gamma_\infty := 2\delta_\infty$, and this completes the proof. $\square$

5. Examples

The following lemma is useful if one wants to verify the assumptions of Theorem 2.1 and Corollary 2.5. It shows how the growth of the characteristic exponent at 0 (resp., at infinity) is related to the existence of moments of the Lévy measure at infinity (resp., at 0).

5.1. **Lemma.** *Let $\psi : \mathbb{R}^d \to \mathbb{C}$ be a continuous negative definite function with Lévy triplet $(\ell, 0, \nu)$, and let f be a Bernstein function with characteristics $(0, \mu)$.*

(i) *If $\mu(dy) \geq c|y|^{-1-\rho}\,dy$ on $B(0,1)$ for some $c > 0$ and $\rho \in (0,1)$, then*

$$\liminf_{\lambda\to\infty} \frac{f(\lambda)}{\lambda^\rho} > 0.$$

(ii) (a) *$f(r) \leq cr^\delta$ for all $r \geq 1$ implies $\int_{(0,1)} r^{\delta+\epsilon}\,\mu(dr) < \infty$ for any $\epsilon > 0$.*
(b) *$f(r) \leq cr^\delta$ for all $r \in [0,1]$ implies $\int_{(1,\infty)} r^{\delta-\epsilon}\,\mu(dr) < \infty$ for any $\epsilon > 0$.*

(iii) (a) *$|\operatorname{Re}\psi(\xi)| \leq c|\xi|^\alpha$ for all $|\xi| \geq 1$ implies $\int_{B(0,1)} |y|^{\alpha+\epsilon}\,\nu(dy) < \infty$ for any $\epsilon > 0$.*
(b) *$|\operatorname{Re}\psi(\xi)| \leq c|\xi|^\alpha$ for all $|\xi| \leq 1$ implies $\int_{B(0,1)^c} |y|^{\alpha-\epsilon}\,\nu(dy) < \infty$ for any $\epsilon > 0$.*

Proof. (i) Fix $\lambda > 1$. As $1 - e^{-\lambda r} \geq 0$ for $r \geq 0$, we have

$$\begin{aligned} f(\lambda) = \int_{(0,\infty)} (1 - e^{-\lambda r})\,\mu(dr) &\geq \int_{(0,\lambda^{-1})} (1 - e^{-\lambda r})\,\mu(dr) \\ &\geq c\int_{(0,\lambda^{-1})} (1 - e^{-\lambda r})\,\frac{dr}{r^{1+\rho}}. \end{aligned}$$

Changing variables according to $s := \lambda r$, we find that the right-hand side equals $c'\lambda^\rho$ for some strictly positive constant c'.

(ii)(a) If $\delta + \epsilon \geq 1$ there is nothing to show since $\int_{(0,1)} r\,\mu(dr) < \infty$. For $\delta + \epsilon \in (0,1)$ we use the formula

$$r^{\delta+\epsilon} = \frac{\delta+\epsilon}{\Gamma(1-\delta-\epsilon)} \int_{(0,\infty)} (1 - e^{-rs})\,\frac{ds}{s^{1+\delta+\epsilon}} \tag{34}$$

It is not difficult to see that this implies

$$r^{\delta+\epsilon} \leq C\int_{(1,\infty)} (1 - e^{-rs})\,\frac{ds}{s^{1+\delta+\epsilon}} \quad \text{for all} \quad r \in (0,1)$$

for some constant $C > (\delta+\epsilon)/\Gamma(1-\delta-\epsilon)$. Applying Tonelli's theorem we find

$$\int_{(0,1)} r^{\delta+\epsilon}\,\mu(dr) \leq C\int_{(1,\infty)}\int_{(0,1)} (1 - e^{-rs})\,\mu(dr)\,\frac{ds}{s^{1+\delta+\epsilon}} \leq C\int_{(1,\infty)} \frac{f(s)}{s^{1+\delta+\epsilon}}\,ds < \infty.$$

(ii)(b) Since f grows at most linearly and $\int_{|y|\geq 1} \mu(dy) < \infty$, we can assume without loss of generality that $\delta \in (0,1]$ and $\delta - \epsilon > 0$. It follows from (34) (with ϵ replaced by $-\epsilon$) that there exists a constant $c' > 0$ such that

$$r^{\delta-\epsilon} \leq c'\int_{(0,1)} (1 - e^{-rs})\,\frac{ds}{s^{1+\delta-\epsilon}} \quad \text{for all} \quad r \geq 1.$$

Applying Tonelli's theorem once again shows

$$\int_{(1,\infty)} r^{\delta-\epsilon}\,\mu(dr) \leq c'\int_{(0,1)} \frac{f(s)}{s^{1+\delta-\epsilon}}\,ds < \infty.$$

(iii) The reasoning is very similar to the proof of (ii); use that for $\alpha \in (0,2)$

$$|\xi|^\alpha = \frac{\alpha 2^{\alpha-1}\Gamma\left(\frac{\alpha+d}{2}\right)}{\pi^{d/2}\,\Gamma\left(1-\frac{\alpha}{2}\right)} \int_{\mathbb{R}^d\setminus\{0\}} (1 - \cos(\xi y))\,\frac{dy}{|y|^{d+\alpha}}, \quad \xi \in \mathbb{R}^d;$$

see also [13, Lemma A.1]. □

Combining Lemma 5.1 with Corollary 2.5 we get the following statement.

5.2. **Example.** Let $(L_t)_{t\geq 0}$ be a d-dimensional Lévy process with one of the following characteristic exponents $\psi : \mathbb{R}^d \to \mathbb{R}$:

(i) $\psi(\xi) = |\xi|^\alpha$ for $\alpha \in (1,2]$; (isotropic stable)

(ii) $\psi(\xi) = (|\xi|^2 + m^2)^{\alpha/2} - m^\alpha$ for $\alpha \in (1,2)$, $m > 0$; (relativistic stable)

(iii) $\psi(\xi) = -(|\xi|^2 + m^2)^{\alpha/2} \cos\left(\alpha \arctan \frac{|\xi|}{m}\right) + m^\alpha$ for $\alpha \in (1,2)$, $m > 0$; (tempered stable)

(iv) $\psi(\xi) = (|\xi|^2 + m)_\alpha - (m)_\alpha$ for some $\alpha \in (1,2)$, $m > 0$; (Lamperti stable) here $(t)_\alpha := \Gamma(t+\alpha)/\Gamma(t)$ denotes the Pochhammer symbol.

If $b : (0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$ is a bounded function which is β-Hölder continuous with respect to x and η-Hölder continuous with respect to t for some $\eta \in (0,1]$ and $\beta > \frac{2}{\alpha} - 1$, then the SDE

$$dX_t = b(t, X_{t-})\,dt + dL_t, \quad X_0 = x \in \mathbb{R}^d,$$

has a pathwise unique strong solution. For any $p < \gamma_\infty$ and $T > 0$ there exists a constant $C > 0$ such that

$$\mathbb{E}\left(\sup_{t\leq T} |X_t - X_t^{(n)}|^p\right) \leq C n^{-\min\{1,\, p\beta/\alpha,\, p\eta\}} \quad \text{for all} \quad n \geq 1$$

where we set $\gamma_\infty := \alpha$ for the exponent (i) and $\gamma_\infty := \infty$ for all other exponents (ii)–(iv).

5.3. **Remark.** (i) The Lévy measure of a tempered stable Lévy process is given by

$$\nu(dy) = \frac{1}{2}\frac{\alpha(\alpha-1)}{\Gamma(2-\alpha)} e^{-m|y|}|y|^{-d-\alpha}\,dy \quad \text{for } \alpha \in (1,2]$$

cf. [9] or [12, Example 5.7]. Note that different authors use different names for this process, e.g. KoBoL process, CGMY process and truncated Lévy process.

(ii) Example 5.2 can be also shown by combining Theorem 2.1 with the heat kernel estimates established in [12], see also [11]; in fact, any continuous negative definite function listed in [12, Table 2] satisfies the assumptions of Theorem 2.1.

We close this section with a further example; it covers many interesting and important Lévy processes.

5.4. **Example.** Let $(L_t)_{t\geq 0}$ be a d-dimensional Lévy process with characteristic exponent ψ and Lévy triplet $(0,0,\nu)$. Assume that ν is of the form

$$\nu(A) = \int_{\mathbb{S}^{d-1}} \int_{(0,\infty)} \mathbb{1}_A(r\vartheta) Q(r)\,dr\,\mu(d\vartheta), \quad A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\}) \tag{35}$$

for a non-trivial measure μ on the unit sphere $\mathbb{S}^{d-1}$ in $\mathbb{R}^d$ and a function $Q : (0,\infty) \to (0,\infty)$ satisfying

$$0 < \liminf_{r\to 0} \frac{Q(r)}{r^{1+\gamma_0}} \leq \limsup_{r\to 0} \frac{Q(r)}{r^{1+\gamma_0}} < \infty \quad \text{and} \quad \limsup_{r\to\infty} \frac{Q(r)}{r^{1+\gamma_\infty}} < \infty$$

for some $\gamma_0 \in (1,2]$ and $\gamma_\infty > 0$. If $b : (0,\infty) \times \mathbb{R}^d \to \mathbb{R}^d$ is a bounded function which is β-Hölder continuous with respect to x and η-Hölder continuous with respect to t for some $\eta \in (0,1]$ and

$$\gamma_0(1+\beta) > 2,$$

then the SDE

$$dX_t = b(t, X_{t-})\,dt + dL_t, \quad X_0 = x \in \mathbb{R}^d,$$

has a pathwise unique strong solution. For any $p < \gamma_\infty$ and $T > 0$ there exists a constant $C > 0$ such that

$$\mathbb{E}\left(\sup_{t\leq T} |X_t - X_t^{(n)}|^p\right) \leq C n^{-\min\{1,\, p\beta/\gamma_0,\, p\eta\}} \quad \text{for all} \quad n \geq 1.$$

Since each of the processes in Example 5.2 has a Lévy measure of the form (35), Example 5.4 is more general than Example 5.2. Let us point out that Example 5.4 includes truncated stable Lévy processes, i.e. $Q(r) = r^{-1-\alpha}\mathbb{1}_{(0,1)}(r)$, and layered stable Lévy processes, i.e. $Q(r) = r^{-1-\alpha}\mathbb{1}_{(0,1)}(r) + r^{-1-\beta}\mathbb{1}_{[1,\infty)}(r)$.

Proof of Example 5.4. Some elementary calculations show that $c^{-1}|\xi|^{\gamma_0} \leq \operatorname{Re}\psi(\xi) \leq c|\xi|^{\gamma_0}$ for some constant $c \in (0,\infty)$ as $|\xi| \to \infty$. Moreover, by the very definition of ν, $\int_{|y|\leq 1} |y|^{\gamma_0+\epsilon}\,\nu(dy) + \int_{|y|\geq 1} |y|^{\gamma_\infty-\epsilon}\,\nu(dy) < \infty$ for any $\epsilon > 0$. Applying Corollary 2.4 with $f(r) := r^{\gamma_0}$ finishes the proof. □

Appendix A.

For the proof of our main results we use the following auxiliary statements.

A.1. **Proposition** (differentiation lemma for parameter-dependent integrals)**.** *Let $(X,\mathcal{A},\mu)$ be a σ-finite measure space and $\phi : (a,b)\times X \to \mathbb{R}$ a measurable function with the following properties.*

(i) $\int_X |\phi(s,x)|\,\mu(dx) < \infty$ *for all* $s \in (a,b)$.

(ii) $s \mapsto \phi(s,x)$ *is differentiable for all* $x \in X$ *and*

$$\int_{(a,b)} \int_X |\partial_s \phi(s,x)|\,\mu(dx)\,ds < \infty.$$

(iii) $s \mapsto \int_X \partial_s\phi(s,x)\,\mu(dx)$ *is continuous.*

Then $F(s) := \int_X \phi(s,x)\,\mu(dx)$ *is continuously differentiable for all* $s \in (a,b)$ *and*

$$F'(s) = \int_X \partial_s\phi(s,x)\,\mu(dx), \quad s \in (a,b).$$

Note that (iii) is always satisfied if $t \mapsto \partial_t\phi(t,x)$ is continuous and there exists a function $w \in L^1(\mu)$ such that $|\partial_t\phi(t,x)| \leq w(x)$ for all $t \in (a,b)$ and $x \in X$; therefore, Proposition A.1 extends the standard version of the differentiation lemma which can be found, for instance, in [24, Theorem 12.5].

Proof of Proposition A.1. Fix $s \in (a,b)$. Applying the fundamental theorem of calculus and Fubini's theorem, we find

$$\begin{aligned} F(s+h) - F(s) = \int_X (\phi(s+h,x) - \phi(s,x))\,\mu(dx) &= \int_X \int_s^{s+h} \partial_r\phi(r,x)\,dr\,\mu(dx) \\ &= \int_s^{s+h} \int_X \partial_r\phi(r,x)\,\mu(dx)\,dr \end{aligned}$$

for all $h \in \mathbb{R}$. By assumption, $f(r) := \int_X \partial_s\phi(r,x)\,\mu(dx)$ is continuous, and so

$$\lim_{h\to 0} \frac{1}{h}(F(s+h) - F(s)) = \lim_{h\to 0} \frac{1}{h}\int_s^{s+h} f(r)\,dr = f(s) \overset{\text{def}}{=} \int_X \partial_s\phi(s,x)\,\mu(dx).$$ □

A.2. **Proposition.** *Let $(L_t)_{t\geq 0}$ be a k-dimensional Lévy process with Lévy triplet $(\ell,0,\nu)$ and jump measure N such that $\int_{|y|\geq 1} |y|^\gamma\,\nu(dy) < \infty$ holds for some $\gamma \in [1,2]$. Denote by $\mathcal{F} := (\mathcal{F}_t)_{t\geq 0}$, $\mathcal{F} := \sigma(L_s; s \leq t)$, the natural filtration and let $b : [0,\infty)\times\Omega \to \mathbb{R}^d$ and $\sigma : [0,\infty)\times\Omega \to \mathbb{R}^{d\times k}$ be $\mathcal{F}$-progressively measurable bounded functions. Then the process*

$$X_t := x + \int_0^t b(s)\,ds + \int_0^t \sigma(s)\,dL_s, \quad x \in \mathbb{R}^d$$

satisfies Itô's formula

$$\begin{aligned} &F(t,X_t) - F(0,X_0) \\ &= \int_{(0,t)} \partial_s F(s,X_s)\,ds + \int_{(0,t)} \nabla_x F(s,X_s)\cdot(b(s) + \sigma(s)\cdot\ell)\,ds \\ &\quad + \iint_{(0,t)\times B(0,1)} (F(s,X_{s-} + \sigma(s)\cdot y) - F(s,X_{s-}))\,\tilde{N}(dy,ds) \\ &\quad + \iint_{(0,t)\times B(0,1)^c} (F(s,X_{s-} + \sigma(s)\cdot y) - F(s,X_{s-}))\,N(dy,ds) \\ &\quad + \iint_{(0,t)\times\mathbb{R}^d} (F(s,X_{s-} + \sigma(s)\cdot y) - F(s,X_{s-}) - \nabla_x F(s,X_{s-})\cdot\sigma(s)y\mathbb{1}_{(0,1)}(|y|))\,\nu(dy)\,ds \end{aligned} \tag{36}$$

for any function $F \in C_b^{1,1}((0,\infty)\times\mathbb{R}^d)$ *such that for every* $T>0$

$$\|\nabla_x F\|_{\mathcal{C}^{\gamma-1}([0,T])} := \sup_{t\in[0,T]}\sup_{x\in\mathbb{R}^d}|\nabla_x F(t,x)| + \sup_{t\in[0,T]}\sup_{\substack{x,y\in\mathbb{R}^d\\ x\neq y}}\frac{|\nabla_x F(t,x)-\nabla_y F(t,y)|}{|x-y|^{\gamma-1}} < \infty.$$

Sketch of the proof. Fix $T>0$ and pick $\chi\in C_c^\infty(\mathbb{R}^d)$ such that $\chi\geq 0$, $\int\chi(y)\,dy=1$ and $\chi(y)=0$ for all $|y|\geq 1$. If we set

$$F_k(t,x) := k^d\int_{\mathbb{R}^d}F(t,x+y)\chi(ky)\,dy,\quad x\in\mathbb{R}^d,\ t\geq 0,\ k\in\mathbb{N},$$

then $F_k\in C_b^{1,2}((0,\infty)\times\mathbb{R}^d)$,

$$\sup_{t\in[0,T]}\sup_{x\in K}\big(|(F-F_k)(t,x)|+|\nabla_x(F-F_k)(t,x)|+|\partial_t(F-F_k)(t,x)|\big)\xrightarrow{k\to\infty}0$$

for any compact set $K\subseteq\mathbb{R}^d$; moreover, we have

$$\sup_{k\in\mathbb{N}}\|\nabla_x F_k\|_{\mathcal{C}^{\gamma-1}([0,T])}\leq\|\nabla_x F\|_{\mathcal{C}^{\gamma-1}([0,T])}<\infty.$$

Applying Taylor's formula we find that there exist a sequence $c_k\to 0$ and a constant $C>0$ such that

$$\begin{aligned}|(F-F_k)(s,x+z)-(F-F_k)(s,x)|&\leq c_k\min\{1,|z|\}\\ |(F-F_k)(s,x+z)-(F-F_k)(s,x)-\nabla_x(F-F_k)(s,x)\cdot z|&\leq C\min\{1,|z|^\gamma\}\end{aligned}\tag{37}$$

for all $x\in K$, $z\in\mathbb{R}^d$ and $s\in[0,T]$. Since we can apply Itô's formula for each $F_k\in C_b^{1,2}((0,\infty)\times\mathbb{R}^d)$, see e.g. [6, Chapter II.5], we get (36) for $F=F_k$. Define

$$\tau_R := \inf\{t\geq 0; |X_t|\geq R\}.$$

Using (36) for F_k and replacing t by $t\wedge\tau_R$, it is not difficult to see that each of the integrals converges as $k\to\infty$: for the third integral (which is an L^2-martingale), we use Itô's isometry, (37) and the dominated convergence theorem to get L^2-convergence and then we extract an almost surely convergent subsequence; all other integral expressions converge almost surely because of (37) and dominated convergence. This gives (36) for F and with t replaced by $t\wedge\tau_R$. Almost the same argument allows us now to let $R\to\infty$, and the claim follows. □

A.3. **Lemma.** *Let* $(X_t^{(n)})_{t\in[0,T]}$ *be a sequence of stochastic processes with càdlàg sample paths such that*

$$\mathbb{E}\left(\sup_{0\leq t\leq T}\big|X_t^{(n)}-X_t^{(m)}\big|^p\right)\xrightarrow{m,n\to\infty}0$$

for some $p>0$. *Then there exists a stochastic process* $(X_t)_{t\in[0,T]}$ *and a subsequence* $(n_k)_{k\in\mathbb{N}}$ *such that* $\sup_{t\in[0,T]}|X_t-X_t^{(n_k)}|\to 0$ *almost surely as* $k\to\infty$.

Proof. For $p\geq 1$ this follows from the Riesz–Fischer theorem on the completeness of the spaces L^p, see e.g. [24, Theorem 13.7]; therefore it suffices to consider the case $p\in(0,1)$. For $k\geq 1$ choose iteratively $n_k>n_{k-1}$ such that

$$\mathbb{E}\left(\sup_{0\leq t\leq T}\big|X_t^{(n_k)}-X_t^{(n_{k-1})}\big|^p\right)\leq\frac{1}{2^k}.$$

As $p\in(0,1)$, we have $(x+y)\leq x^p+y^p$ for $x,y\geq 0$, and this implies

$$\mathbb{E}\left(\sum_{k\geq 1}\sup_{0\leq t\leq T}\big|X_t^{(n_k)}-X_t^{(n_{k-1})}\big|^p\right)\leq\sum_{k\geq 1}\mathbb{E}\left(\sup_{0\leq t\leq T}\big|X_t^{(n_k)}-X_t^{(n_{k-1})}\big|^p\right)\leq\sum_{k\geq 1}\frac{1}{2^k}<\infty.$$

Since $X_t^{(n_i)}=X_t^{(n_0)}+\sum_{k=1}^i(X_t^{(n_k)}-X_t^{(n_{k-1})})$ this shows that the limit $X_t:=\lim_{i\to\infty}X_t^{(n_i)}$ exists uniformly in $t\in[0,T]$ with probability 1. □

References

[1] Chen, Z.-Q., Song, R., Zhang, X.: Stochastic flows for Lévy processes with Hölder drifts. Preprint arXiv 1501.04758.

[2] Deng, C.-S., Schilling, R. L.: On shift Harnack inequalities for subordinate semigroups and moment estimates for Lévy processes. *Stoch. Proc. Appl.* **125** (2015), 3851–3878.

[3] Haadem, S., Proske, F.: On the Construction and Malliavin Differentiability of Lévy Noise Driven SDEs with Singular Coefficients. *J. Funct. Anal.* **266** (2014), 5321–5359.

[4] Hashimoto, H.: Approximation and stability of solutions of SDEs driven by a symmetric α stable process with non-Lipschitz coefficients. In: Donati, M. C., Leiay, A., Rouault, A. (eds): *Séminaire de Probabilités XLV*. Lecture Notes in Mathematics **2078**. Springer, 2013.

[5] Higham, D. J, Kloeden, P. E.: Strong convergence rates for backward Euler on a class of non-linear jump-diffusion problems. *J. Computat. Appl. Math.* **205** (2007), 949–956.

[6] Ikeda, N., Watanabe, S.: *Stochastic Differential Equations and Diffusion Processes*. North-Holland, Amsterdam 1989 (2nd edn).

[7] Jacod, J.: The Euler scheme for Lévy driven stochastic differential equations: Limit theorems. *Ann. Probab.* **32** (2004), 1830–1872.

[8] Knopova, V., Schilling, R. L.: A note on the existence of transition probability densities for Lévy processes. *Forum Math.* **25** (2013), 125–149.

[9] Küchler, U., Tappe, S.: Tempered stable distributions and processes. *Stoch. Proc. Appl.* **123** (2013), 4256–4293.

[10] Kühn, F.: Existence and estimates of moments for Lévy-type processes. *Stoch. Proc. Appl.* **127** (2017), 1018–1041.

[11] Kühn, F.: Transition probabilities of Lévy-type processes: Parametrix construction. Preprint arXiv 1702.00778.

[12] Kühn, F.: *Lévy-Type Processes: Moments, Construction and Heat Kernel Estimates*. Springer Lecture Notes in Mathematics **2187** (vol. VI of the "Lévy Matters" subseries). Springer, to appear 2017.

[13] Kühn, F., Schilling, R. L.: On the domain of fractional Laplacians and related generators of Feller processes. Preprint arXiv 1610.08197.

[14] Kühn, F., Schilling, R. L.: A probabilistic proof of Schoenberg's theorem. *In preparation.*

[15] Kulik, A.: On weak uniqueness and distributional properties of a solution to an SDE with α-stable noise. Preprint arXiv 1511.00106.

[16] Lunardi, A.: *Interpolation Theory*. Edizioni Della Normale, Pisa 2009 (2nd edn).

[17] Matheron, G.: *Les variables régoionalisées et leur estimation. Une application de la théorie des fonctions aléatoires aux sciences de la nature*. Masson & Cie., Paris 1965.

[18] Mikulevicus, R., Xu, F.: On the rate of convergence of strong Euler approximation for SDEs driven by Lévy processes. Preprint arXiv 1608.02303.

[19] Novikov, A. A.: On discontinuous martingales. *Theory Probab. Appl.* **20** (1975), 11–26.

[20] Pamen, O. M., Taguchi, D.: Strong rate of convergence for the Euler–Maruyama approximation of SDE with Hölder continuous drift coefficient. *Stoch. Proc. Appl.* **127** (2017), 2542–2559.

[21] Protter, P.: *Stochastic Integration and Differential Equations*. Springer, Berlin 2004 (2nd edn).

[22] Qiao, H.: Euler–Maruyama approximation for SDEs with jumps and non-Lipschitz coefficients. *Osaka J. Math.* **51** (2014), 47–66.

[23] Sato, K.: *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge 2005.

[24] Schilling, R. L.: *Measures, Integrals and Martingales*. Cambridge University Press, Cambridge 2017 (2nd edn).

[25] Schilling, R. L., Song, R., Vondraček, Z.: *Bernstein functions*. De Gruyter, Berlin 2012 (2nd edn).

[26] Schilling, R. L., Sztonyk, P., Wang, J.: Coupling property and gradient estimates of Lévy processes via the symbol. *Bernoulli* **18** (2012), 1128–1149.

[27] Tanaka, H., Tsuchiya, M., Watanabe, S.: Perturbation of drift-type for Lévy processes. *J. Math. Kyoto Univ.* **14** (1974), 73–92.

[28] Triebel, H.: *Interpolation Theory, Function Spaces, Differential Operators*. Johann Ambrosius Barth, Heidelberg 1995 (2nd edn).

(F. Kühn) TU Dresden, Fachrichtung Mathematik, Institut für Mathematische Stochastik, 01062 Dresden, Germany

E-mail address: `franziska.kuehn1@tu-dresden.de`

(R. L. Schilling) TU Dresden, Fachrichtung Mathematik, Institut für Mathematische Stochastik, 01062 Dresden, Germany

E-mail address: `rene.schilling@tu-dresden.de`