



Deposited via The University of Leeds.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/236364/>

Version: Accepted Version

Article:

Butkovsky, O., Dareiotis, K. and Gerencsér, M. (2025) Strong rate of convergence of the Euler scheme for SDEs with irregular drift driven by Lévy noise. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques*, 61 (4). pp. 2624-2660. ISSN: 0246-0203 (In Press)

<https://doi.org/10.1214/24-aihp1506>

Copyright © 2025 Association des Publications de l'Institut Henri Poincaré. This is an author produced version of an article published in *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*. Uploaded in accordance with the publisher's self-archiving policy.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Strong rate of convergence of the Euler scheme for SDEs with irregular drift driven by Lévy noise

January 12, 2024

Oleg Butkovsky*, Konstantinos Dareiotis[†], and Máté Gerencsér[‡]

Abstract

We study the strong rate of convergence of the Euler–Maruyama scheme for a multidimensional stochastic differential equation (SDE)

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

with irregular β -Hölder drift, $\beta > 0$, driven by a Lévy process with exponent $\alpha \in (0, 2]$. For $\alpha \in [2/3, 2]$, we obtain strong L_p and almost sure convergence rates in the entire range $\beta > 1 - \alpha/2$, where the SDE is known to be strongly well-posed. This significantly improves the current state of the art, both in terms of convergence rate and the range of α . Notably, the obtained convergence rate does not depend on p , which is a novelty even in the case of smooth drifts. As a corollary of the obtained moment-independent error rate, we show that the Euler–Maruyama scheme for such SDEs converges almost surely and obtain an explicit convergence rate. Additionally, as a byproduct of our results, we derive strong L_p convergence rates for approximations of nonsmooth additive functionals of a Lévy process. Our technique is based on a new extension of stochastic sewing arguments and Le’s quantitative John–Nirenberg inequality.

1. Introduction

We consider the stochastic differential equation

$$dX_t = b(X_t) dt + dL_t, \quad t \geq 0, \quad X_0 = x_0, \quad (1.1)$$

driven by a d -dimensional Lévy process L . Here the coefficient b is a measurable function $\mathbb{R}^d \rightarrow \mathbb{R}^d$, and the initial condition $x_0 \in \mathbb{R}^d$. Throughout the article the dimension $d \in \mathbb{N} = \{1, 2, \dots\}$ is arbitrary.

The ‘strength’ of a Lévy process can often be characterised by a single parameter $\alpha \in (0, 2]$ called the *stable index* (for various examples see Section 2.1), with $\alpha = 2$ corresponding to the usual Brownian motion. This parameter can be used to describe the regularisation provided by L . Indeed, assuming a natural nondegeneracy condition on the jump measure of L , (1.1) has a unique strong solution whenever b belongs to the Hölder space $\mathcal{C}^\beta(\mathbb{R}^d)$, where β satisfies

$$\beta > 1 - \frac{\alpha}{2}. \quad (1.2)$$

*Weierstrass Institute, Mohrenstraße 39, 10117 Berlin, Germany oleg.butkovskiy@gmail.com

[†]University of Leeds, Woodhouse, LS2 9JT Leeds, United Kingdom k.dareiotis@leeds.ac.uk

[‡]TU Wien, Wiedner Hauptstraße 8-10, 1040 Vienna, Austria mate.gerencser@tuwien.ac.at

This solution theory was developed in [Pri12, Pri15, Pri18, CSZ18].

Once the well-posedness is understood, it is natural to investigate basic discretisations of the equation. The most classical Euler-Maruyama approximation of (1.1) reads as

$$dX_t^n = b(X_{\kappa_n(t)}^n) dt + dL_t, \quad X_0^n = x_0^n, \quad (1.3)$$

with $\kappa_n(t) := \lfloor nt \rfloor n^{-1}$. For Lipschitz b the convergence of X^n to X can be seen by very straightforward arguments. The L_p -rate of convergence obtained this way, however, deteriorates for large p (see Remark 2.7 (ii) below for some details). Even though our main focus will be on non-Lipschitz coefficients, it is noteworthy and perhaps surprising that our method coincidentally also solves this moment issue and therefore gives new results even in the Lipschitz case.

One of the first results regarding convergence of X^n to X in the case of irregular b is [PT17]. It is shown there that

$$\left\| \sup_{t \in [0,1]} |X_t - X_t^n| \right\|_{L_p(\Omega)} \leq N n^{-(\frac{\beta}{2} \wedge \frac{1}{p})}, \quad n \in \mathbb{N}, \quad (1.4)$$

for the case where L is a truncated symmetric α -stable process, $\alpha \in (1, 2)$, $\beta > 2 - \alpha$; here $N = N(\alpha, \beta, p, d)$ is a certain positive constant. This result was improved in [KS19b] in three directions: first, the condition on β is relaxed to $\beta > 2/\alpha - 1$; second, the rate of convergence in (1.4) is improved and is $\frac{\beta}{\alpha} \wedge \frac{1}{p}$; third, the class of considered Lévy processes is significantly extended and additionally includes standard isotropic stable processes, tempered stable processes, relativistic stable processes, and others. The same rate $\frac{\beta}{\alpha} \wedge \frac{1}{p}$ is derived [MX18] even for multiplicative noise, under the further relaxed condition $\beta > 1 - \alpha/2$, under which strong solutions are known to exist. Very recently, for the case of standard isotropic stable processes, [HSY21] showed that the strong L_p rate of $\frac{\beta}{\alpha} \wedge \frac{1}{p}$ holds in the whole range $\beta > 1 - \alpha/2$ even when β denotes regularity only in a certain Sobolev scale. A standard example of a coefficient that possesses Sobolev but not Hölder regularity is one with discontinuity of the first kind, in this case scalar SDEs driven by a Brownian motion and a finite activity Poisson process were studied in [PS21].

From the discussion above, the reader may notice the following gaps in the literature. All of the works mentioned above consider the case $\alpha \in [1, 2]$; recall however that the strong well-posedness of (1.1) is known in the whole range $\alpha \in (0, 2]$. Further, the aforementioned moment issue is still present: the convergence rate becomes arbitrary slow for very large p . Note that this also has the consequence that one can not deduce almost sure convergence of $|X_t - X_t^n|$ as $n \rightarrow \infty$. Indeed, to show this one has to prove the bound $\mathbb{E}|X_t - X_t^n|^p \leq N n^{-1-\varepsilon}$ for some $p > 0$, $\varepsilon > 0$, while the best available bound is $\mathbb{E}|X_t - X_t^n|^p \leq N n^{-1}$, which is not sufficient.

The present paper closes these gaps; the novelties can be summarised as follows. First, our methods are completely free from the moment issue. As alluded to above and detailed in Remark 2.7 (ii) below, this makes our results new even in the smooth drifts, where the strong L_p rate, which one trivially gets, namely, $\frac{1}{\alpha} \wedge \frac{1}{p}$, $p \geq 1$, is improved to 1. Another factor that may provide poor rate in the previous results is if β/α is small. This issue is also not present here, the obtained rate is always strictly above $1/2$. We also obtain almost sure convergence (and rate) of the Euler-Maruyama scheme, to our knowledge for the first time for SDEs of the type (1.1).

Second, we are not restricted to $\alpha \in [1, 2]$. In the regime $\alpha \in [2/3, 2]$ our assumption on β coincides with the optimal condition (1.2). We cover some (but not the optimal)

range of irregular drifts when $\alpha \in (1/2, 2/3)$. For $\alpha \in (0, 1/2]$ we require b to be regular (i.e. more than Lipschitz), the contribution in this case is the handling of high moments.

Third, we have also shown that the solution of (1.1) is the limit of the corresponding Picard's successive approximations. This extends to the non-smooth case the corresponding results from [Yam81, Tan92, NO23].

Fourth, the class of considered driving Lévy processes is fairly large. It is similar to [KS19b], so it includes not only the standard stable processes but also their different “relatives”. The conditions on the Lévy process, see H1–H3 below, are rather general and are easy to verify for various examples, see Section 2.1.

Finally, as a byproduct of our proofs, we obtain significant improvements in a different but related problem of numerical stochastic analysis: the question of approximating additive functionals of stochastic processes given high-frequency observations. To wit, consider an \mathbb{R}^d -valued stochastic process Y , a measurable function $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$, and the ‘occupation time functional’

$$\Gamma = \int_0^1 f(Y_s) ds. \quad (1.5)$$

Approximation of Γ is an interesting question in itself (for a detailed overview of the literature, we refer to [Alt21, Section 1]). A natural estimation scheme is given by

$$\Gamma^n = \int_0^1 f(Y_{\kappa_n(s)}^n) ds = \frac{1}{n} \sum_{i=0}^n f(Y_{i/n}^n), \quad (1.6)$$

where Y^n is the process approximating Y . If the increments of Y can be simulated directly, then one can, of course, take $Y^n := Y$. For the case where Y is a Markov process, whose density satisfies certain estimates (e.g., Y is an α -stable process), $Y^n = Y$, and f is a bounded or Hölder continuous function, L_p error bounds for $\Gamma - \Gamma^n$ were obtained in [GK14, GKK15]. For the case $p = 2$, Altmeyer [Alt21] improved the convergence rate to the one which was shown to be optimal in some setups [AL22, AJP22].

In this article, we take the best of both worlds. Namely, we obtain the L_p convergence rate for the error $\Gamma - \Gamma^n$ as in [Alt21] but for general $p \geq 2$. Furthermore, we provide L_p bounds on the convergence rate of the error for the case $Y^n \neq Y$. This is relevant if the process Y cannot be simulated directly (this is often the case if Y is a solution to SDE (1.1)) and we have access only to its approximation Y^n (which can be its Euler approximation (1.3)).

Our approach is rather different from the one used in the papers [PT17, MX18, KS19b, HSY21] discussed above, as we do not rely on any form of Zvonkin transformation or Itô-Tanaka trick (in fact, Itô’s formula is not even once applied for either X or X^n). Instead, we employ stochastic sewing, which originates from the work of Lê [Lê20] and has been developed for discretisation problems in [BDG21a, BDG21b, DGL21, LL21]. In [BDG21a] the Euler-Maruyama scheme for fractional Brownian motion-driven SDEs is studied. One may hope due to the scaling correspondance between stable indices and Hurst parameters ($\alpha \leftrightarrow 1/H$) that the methods therein translate easily to the Lévy case. This is unfortunately not the case, for several reasons. First, the usual regime $H < 1$ corresponds to $\alpha > 1$. To consider $\alpha < 1$, one needs tools from the $H > 1$ case, in particular the shifted stochastic sewing lemma [Ger20]. Second, high moments of Lévy processes do not scale (or they do not even exist for several examples). This is related to the aforementioned moment issue yielding poor rate for large moments in preceding

literature. Overcoming this challenge relies on the recently obtained quantitative John-Nirenberg inequality [Lê22a]. Finally, there is no useful form of Girsanov theorem to remove the drift at any part of the analysis.

This article deals only with the strong rate of convergence. Clearly, the weak rate of convergence is at least as good as the strong rate of convergence; already with this simple observation our results imply weak rates which are better than the ones available in the literature [KP02, MZ11] in the range of parameters that we cover. An interesting and challenging question is whether these weak rates can be improved further, and whether the range of β can be upgraded to $\beta > 1 - \alpha$ (in this range SDE (1.1) is weakly well-posed [Kul19, LZ22]). We leave this for the future work.

The rest of the paper is organized as follows. Our main result concerning the L_p convergence of the numerical scheme is formulated in Section 2. Examples of Lévy processes satisfying the assumptions of the convergence theorem are given in Section 2.1. A number of technical tools needed for the proofs are collected in Section 3. The main results are proved in Section 4, whilst the proofs of some technical auxiliary statements are placed in the Appendix.

Acknowledgements. The authors are grateful to Randolph Altmeyer and Mark Podolskij for very helpful discussions regarding approximation of additive functionals of a Lévy process, and to Khoa Lê for many useful conversations and for bringing to our attention the articles [GK14, GKK15]. We would like to thank the referees for their helpful comments and feedback. OB has received funding from the DFG Research Unit FOR 2402 and is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy — The Berlin Mathematics Research Center MATH+ (EXC-2046/1, project ID: 390685689, sub-project EF1-22). MG was funded by the Austrian Science Fund (FWF) Stand-Alone programme P 34992. The project has been conceived during the stay of the authors at the Hausdorff Research Institute for Mathematics (HIM), Bonn. Significant progress on the project has been achieved during the visits of the authors to TU Wien, Mathematisches Forschungsinstitut Oberwolfach (mini-workshop 2207c), and Università degli Studi di Torino. We would like to thank all these institutions and their staff for providing excellent working conditions, support, and hospitality.

2. Main results

We begin by introducing the basic notation. For $\beta \in (0, 1)$ and a Borel subset Q of \mathbb{R}^k , $k \in \mathbb{N}$, let $\mathcal{C}^\beta(Q)$ be the corresponding Hölder space, that is, the set of functions $f: Q \rightarrow \mathbb{R}$ such that

$$\|f\|_{\mathcal{C}^\beta(Q)} := \sup_{x \in Q} |f(x)| + [f]_{\mathcal{C}^\beta(Q)} := \sup_{x \in Q} |f(x)| + \sup_{x \neq y \in Q} \frac{|f(x) - f(y)|}{|x - y|^\beta} < \infty.$$

With a slightly unconventional notation we set $\mathcal{C}^0(Q)$ to be the set of bounded measurable functions (not necessarily continuous) equipped with the supremum norm. The definition of the analogous spaces for \mathbb{R}^d -valued functions is simply understood coordinate-wise. For $\beta \in [1, \infty)$ we denote by $\mathcal{C}^\beta(Q)$ the set of functions whose weak derivatives of order $0, 1, \dots, \lfloor \beta \rfloor$ all have representatives belonging to $\mathcal{C}^{\beta - \lfloor \beta \rfloor}(Q)$. In the particular case $Q = \mathbb{R}^d$

sometimes we will use a shorthand and write \mathcal{C}^β instead of $\mathcal{C}^\beta(\mathbb{R}^d)$ in order not to overcrowd the notation.

Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and on it a d -dimensional Lévy process L , equipped with a right-continuous, complete filtration $\mathbb{F} = (\mathcal{F}_t)_{t \in [0,1]}$. The conditional expectation given \mathcal{F}_t will be denoted by \mathbb{E}^t . The Markov transition semigroup associated to the process L is denoted by $\mathcal{P} = (\mathcal{P}_t)_{t \geq 0}$, and the generator of \mathcal{P} is denoted by \mathcal{L} .

We fix $\alpha \in (0, 2]$, $d \in \mathbb{N}$ and impose the following assumptions on \mathcal{P} and \mathcal{L} .

Assumption H1 (gradient type bound on the semigroup). There exists a constant M such that for any $f \in \mathcal{C}^0(\mathbb{R}^d)$, one has

$$\|\nabla \mathcal{P}_t f\|_{\mathcal{C}^0(\mathbb{R}^d)} \leq M t^{-1/\alpha} \|f\|_{\mathcal{C}^0(\mathbb{R}^d)}, \quad 0 < t \leq 1. \quad (2.1)$$

Assumption H2 (action of the generator). For $\delta = 0, 1$, and any $\varepsilon > 0$ there exists a constant $M = M(\delta, \varepsilon)$ such that for any $f \in \mathcal{C}^{\alpha+\delta+\varepsilon}(\mathbb{R}^d)$ vanishing at infinity, one has

$$\|\mathcal{L}f\|_{\mathcal{C}^\delta(\mathbb{R}^d)} \leq M \|f\|_{\mathcal{C}^{\alpha+\delta+\varepsilon}(\mathbb{R}^d)}. \quad (2.2)$$

Assumption H3 (moment conditions). For any $p \in (0, \alpha)$, $\varepsilon > 0$, there exists a constant $M = M(p, \varepsilon)$ such that

$$\mathbb{E}[|L_t|^p \wedge 1] \leq M t^{\frac{p}{\alpha}-\varepsilon}, \quad 0 < t \leq 1.$$

With some abuse of notation, in the sequel when we refer to the parameter M given a process L satisfying the above assumptions, we understand the collection of all of the M -s in H1-H3.

We provide a long list of examples of processes satisfying H1-H3 in Section 2.1, let us here just briefly mention three of the most standard examples. First let $\alpha \in (0, 2)$ and L be the standard d -dimensional α -stable process. That is, L is a Lévy process whose characteristic function is

$$\mathbb{E}e^{i\langle \lambda, L_t \rangle} = e^{-tc_\alpha |\lambda|^\alpha}, \quad \lambda \in \mathbb{R}^d, t \geq 0,$$

for some constant $c_\alpha > 0$. In this case H1-H3 are satisfied. A similar example is the d -dimensional cylindrical α -stable process, that is, a process L whose coordinates are d independent 1-dimensional standard α -stable processes. Its characteristic function is given by

$$\mathbb{E}e^{i\langle \lambda, L_t \rangle} = e^{-t\tilde{c}_\alpha \sum_{i=1}^d |\lambda_i|^\alpha}, \quad \lambda \in \mathbb{R}^d, t \geq 0,$$

for some constant $\tilde{c}_\alpha > 0$. In this case H1-H3 are also satisfied. Finally the most standard of the most standard examples is the d -dimensional Brownian motion, which satisfies H1-H3 with $\alpha = 2$.

Convention on the operator \wedge . The expression of the form $c_1 + c_2 \wedge c_3$, where $c_i \in \mathbb{R}$, will quite often appear in the paper. We will always mean that in this expression the minimum is taken first and then the addition, thus it equals to $c_1 + (c_2 \wedge c_3) = c_1 + \min(c_2, c_3)$.

Convention on constants. Throughout the paper N denotes a positive constant whose value may change from line to line; its dependence is always specified in the corresponding statement.

We begin with the well-posedness of (1.1). As mentioned in the introduction, this is essentially known, but since none of the available results cover the whole range of exponents and generality of driving processes considered herein, we provide a short proof in the appendix. This is done for the sake of presentation as well as to highlight the usefulness of stochastic sewing for well-posedness of SDEs with jumps; our method for obtaining strong well-posedness is very different from [Pri12, Pri15, Pri18, CSZ18]. Another advantage of our proof strategy is that we get “for free” that the solution of (1.1) is the limit of Picard’s successive approximations. While this fact is known for Lipschitz or essentially Lipschitz drifts [Yam81, Theorem 1], [Tan92, Theorem 2], [NO23, Lemma 6.2 and Theorem 6.1], to the best of our knowledge, this result is new for Hölder continuous drifts.

Define successively a sequence of approximations $Y^{(0)}(t) := \eta + L_t$, $t \in [0, 1]$,

$$Y^{(n+1)}(t) = \eta + \int_0^t b(Y^{(n)}(s)) ds + L_t, \quad t \in [0, 1], \quad n \in \mathbb{Z}_+.$$

Theorem 2.1. *Suppose that L satisfies H1–H3. Let η be a \mathcal{F}_0 –measurable random vector taking values in \mathbb{R}^d . Suppose additionally that*

$$\beta > \left(1 - \frac{\alpha}{2}\right) \vee (2 - 2\alpha) \quad (2.3)$$

and let $b \in \mathcal{C}^\beta(\mathbb{R}^d, \mathbb{R}^d)$. Then equation (1.1) with the initial condition $X_0 = \eta$ has a unique strong solution. Furthermore, this solution X is the limit of the Picard iterations, namely

$$\|X - Y^{(n)}\|_{\mathcal{C}^0([0,1])} \rightarrow 0 \quad \text{as } n \rightarrow \infty \quad \text{a.s. and in } L_p(\Omega), \quad p \geq 1.$$

Now we are ready to present our main results: the L_p convergence of the Euler–Maruyama scheme with an explicit rate. In the statements below, X is a solution to (1.1) with the initial condition $x_0 \in \mathbb{R}^d$, and X^n , $n \in \mathbb{Z}_+$, is its Euler approximation with the initial condition $x_0^n \in \mathbb{R}^d$, which solves (1.3).

Theorem 2.2. *Suppose that L satisfies H1–H3 and that (2.3) holds. Let $b \in \mathcal{C}^\beta(\mathbb{R}^d, \mathbb{R}^d)$, $p > 2$, $\varepsilon > 0$. Then there exists a constant $N = N(d, \alpha, \beta, p, \varepsilon, M, \|b\|_{\mathcal{C}^\beta})$ such that for all $n \in \mathbb{N}$ the following bound holds:*

$$\left\| \|X - X^n\|_{\mathcal{C}^0([0,1])} \right\|_{L_p(\Omega)} \leq N n^{-(\frac{1}{2} + \frac{\beta}{\alpha} \wedge \frac{1}{2}) + \varepsilon} + N |x_0 - x_0^n|. \quad (2.4)$$

One can also show the almost sure convergence of X^n to X .

Corollary 2.3. *Suppose that L satisfies H1–H3 and β satisfies (2.3). Let $b \in \mathcal{C}^\beta(\mathbb{R}^d, \mathbb{R}^d)$. Take $x_0^n = x_0$ for all $n \in \mathbb{N}$. Then for any $\varepsilon > 0$, there exists an a.s. finite random variable η such that for any $n \in \mathbb{N}$, $\omega \in \Omega$*

$$\|X(\omega) - X^n(\omega)\|_{\mathcal{C}^0([0,1])} \leq \eta(\omega) n^{-(\frac{1}{2} + \frac{\beta}{\alpha} \wedge \frac{1}{2}) + \varepsilon}.$$

The next theorem gives strong L_p rate of convergence of approximations of additive functionals of a Lévy process.

Theorem 2.4. *Suppose that L satisfies H1–H3. Let $p \in [1, \infty)$, $\varepsilon > 0$, $\theta \in [0, 1]$, $f \in \mathcal{C}^\theta(\mathbb{R}^d, \mathbb{R}^d)$. Then there exists a constant $N = N(d, \alpha, p, \theta, \varepsilon, M)$ such that for all $n \in \mathbb{N}$ the following bound holds:*

$$\left\| \sup_{t \in [0, 1]} \left| \int_0^t f(L_r) dr - \int_0^t f(L_{\kappa_n(r)}) dr \right| \right\|_{L_p(\Omega)} \leq N \|f\|_{\mathcal{C}^\theta} n^{-(\frac{1}{2} + \frac{\theta}{\alpha} \wedge \frac{1}{2}) + \varepsilon}. \quad (2.5)$$

In the case $p = 2$ this rate is proven in [Alt21, Theorem 11]. For general p , a worse rate $\frac{1}{2} + \frac{\theta}{2\alpha} \wedge \frac{1}{2} - \varepsilon$ is proved in [GK14, Theorem 2.2, Section 2.4]. For special cases of f (indicator function or Dirac- δ) even central limit theorems are available in the literature, see e.g. [AJP22, Theorem 2.3], and also [AL22].

The next theorem concerns approximations of additive functionals of the solution X of an SDE driven by a Lévy process. In this case the underlying process X can not be simulated directly, and one has to use the corresponding Euler approximation X^n . We were not able to find any relevant results in this direction in the literature.

Theorem 2.5. *Suppose that L satisfies H1–H3 and that (2.3) holds. Let $b \in \mathcal{C}^\beta(\mathbb{R}^d, \mathbb{R}^d)$, $f \in \mathcal{C}^\theta(\mathbb{R}^d, \mathbb{R}^d)$, where*

$$\theta > \left(1 - \frac{\alpha}{2}\right) \vee \left(2 - 2\alpha\right). \quad (2.6)$$

Let $p \in [1, \infty)$, $\varepsilon > 0$. Then there exists a constant $N = N(d, \alpha, \beta, \theta, p, \varepsilon, M, \|b\|_{\mathcal{C}^\beta})$ such that for all $n \in \mathbb{N}$ the following bound holds:

$$\left\| \sup_{t \in [0, 1]} \left| \int_0^t f(X_r) dr - \int_0^t f(X_{\kappa_n(r)}^n) dr \right| \right\|_{L_p(\Omega)} \leq N \|f\|_{\mathcal{C}^\theta} n^{-(\frac{1}{2} + \frac{\beta \wedge \theta}{\alpha} \wedge \frac{1}{2}) + \varepsilon} + N \|f\|_{\mathcal{C}^\theta} |x_0 - x_0^n|. \quad (2.7)$$

The proofs of these statements are given in Section 4.

Remark 2.6. For any $\kappa \in \mathbb{R}^d$, equations (1.1) and (1.3) can be rewritten as:

$$dX_t = \tilde{b}(X_t) dt + d\tilde{L}_t; \quad dX_t^n = \tilde{b}(X_{\kappa_n(t)}^n) dt + d\tilde{L}_t,$$

where $\tilde{b}(x) := b(x) - \kappa$, $\tilde{L}_t := L_t + \kappa t$. Clearly, if $b \in \mathcal{C}^\beta$, then $\tilde{b} \in \mathcal{C}^\beta$. Therefore, if for some $\kappa \in \mathbb{R}^d$ the process $(L_t + \kappa t)_{t \in [0, 1]}$ satisfies H1–H3, then Theorems 2.2 and 2.5 and Corollary 2.3 hold (provided that all the other conditions on β , b , p are met).

Remark 2.7. Figure 1 shows the region where Theorem 2.2 guarantees the convergence of the Euler scheme. Let us give some context for some of the different regimes of interest for the exponents.

- (i) In the regime $\alpha \in [2/3, 2]$, (2.3) coincides with the well-known condition (1.2), and thus Theorem 2.2 establishes strong convergence in the optimal range of β . We also remark that the threshold $2/3$ appears in the theory of Lévy driven SDEs from time to time, e.g., in [CSZ18, FKM21]. We are unsure whether there is some connection between these appearances, or if this is just an instance of the “law of small numbers”.
- (ii) The result is new even in the case of smooth drift. In the regime $\beta \geq 1$ the drift is regular enough to make the strong well-posedness of (1.1) trivial. Furthermore, it

is easy to get *some* rate of convergence from elementary arguments: assuming for simplicity $x_0 = x_0^n$, by Gronwall's lemma one has almost surely

$$\sup_{t \in [0,1]} |X_t - X_t^n| \leq e^{\|b\|_{C^1}} \|b\|_{C^1} \int_0^1 (|X_t^n - X_{\kappa_n(t)}^n| \wedge 1) dt.$$

This yields for any $p \geq 1$, $\varepsilon > 0$,

$$\left\| \sup_{t \in [0,1]} |X_t - X_t^n| \right\|_{L_p(\Omega)} \leq N n^{-(\frac{1}{\alpha} \wedge \frac{1}{p}) + \varepsilon}, \quad (2.8)$$

with some constant $N = N(\alpha, p, \varepsilon, M, \|b\|_{C^1})$. This provides little control for high moments of the error; further this does not allow to get almost sure rate of convergence. This is markedly improved by Theorem 2.2 and Corollary 2.3.

- (iii) Whenever $\beta \geq \alpha/2$ (which is enforced by (2.3) for all $\alpha \leq 1$), the minimum in $\frac{\beta}{\alpha} \wedge \frac{1}{2}$ is the second term, and so in this case the expression for the L_p -rate simplifies to $1 - \varepsilon$. Thus in this regime we recover the best possible (up to ε) rate for an Euler-type approximation.

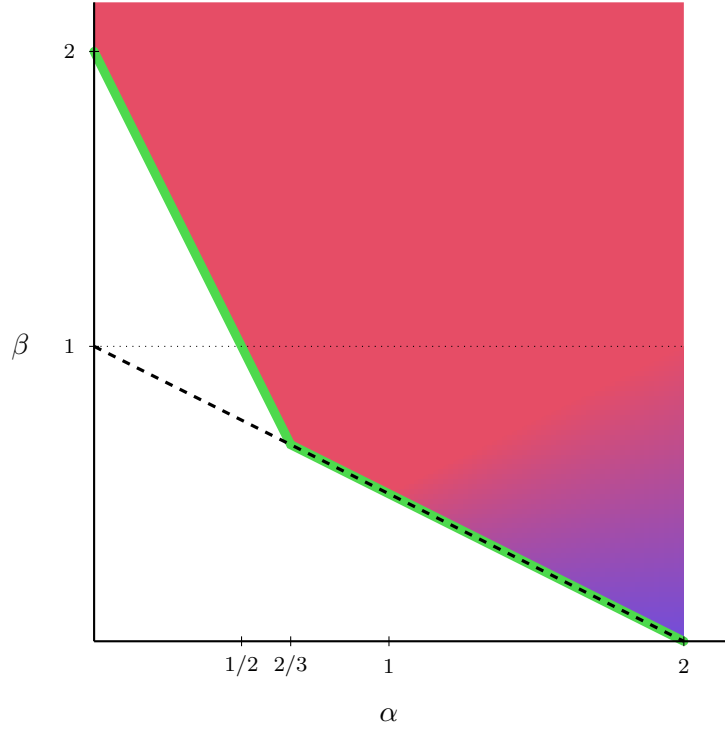


Figure 1: Convergence rates. — is the required lower bound for β ; ---- is the classical condition (1.2); shading indicates rate of L_p -convergence from $1/2$ (blue) to 1 (red).

In the case when L is a standard Brownian motion, Theorem 2.2 is consistent with the results of [BDG21a] in the case $H = 1/2$.

An implementation of the Euler–Maruyama scheme (1.3) would also require simulation of the driving Lévy process L . This is a well-studied problem, and many techniques and tricks are available in the literature. We do not discuss this problem here but rather refer the reader to a concise collection of methods for simulating L provided in [Pap08, Section 14].

2.1 Examples of Lévy processes satisfying the main assumptions

Let us present some examples of Lévy processes satisfying H1–H3. Denote by Φ the characteristic exponent (symbol) of L , that is,

$$\mathbb{E}e^{i\langle \lambda, L_t \rangle} = e^{-t\Phi(\lambda)}, \quad \lambda \in \mathbb{R}^d, t \geq 0.$$

Recall that (see, e.g., [App09, Corollary 2.4.20]) Φ can be written in the form

$$\Phi(\lambda) = -i\langle a, \lambda \rangle + \frac{1}{2}\langle \lambda, Q\lambda \rangle + \int_{\mathbb{R}^d} (1 - e^{i\langle \lambda, y \rangle} + i\langle \lambda, y \rangle \mathbb{1}_{|y| \leq 1}) \nu(dy), \quad \lambda \in \mathbb{R}^d,$$

where $a \in \mathbb{R}^d$, $Q \in \mathbb{R}^{d \times d}$ is a positive semidefinite matrix, and ν is a σ -finite measure on \mathbb{R}^d such that $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}^d} (1 \wedge |y|^2) \nu(dy) < \infty$. It is common to refer to (a, Q, ν) as the generating triplet of L . We begin with general sufficient conditions on Φ , and then move on to the specific examples. By $\operatorname{Re} \Phi$ we mean the real part of Φ .

Proposition 2.8. *Assume that for some $\alpha \in (0, 2)$, $c_1, c_2, N > 0$ the symbol Φ satisfies*

$$c_1|\lambda|^\alpha \leq \operatorname{Re} \Phi(\lambda) \leq c_2|\lambda|^\alpha, \quad \text{when } |\lambda| > N. \quad (2.9)$$

Then the following hold:

- (i) *if $\alpha \in [1, 2)$, then H1 and H2 are satisfied for the process L ;*
- (ii) *if $\alpha \in (0, 1)$, then $\int_{|y| \leq 1} |y| \nu(dy) < \infty$ and H1 and H2 are satisfied for the process $\tilde{L}_t := L_t + \kappa t$, where $\kappa = -a + \int_{|y| \leq 1} y \nu(dy)$.*

Recall that thanks to Remark 2.6 it is sufficient in Theorem 2.2 to verify Assumptions H1 and H2 for the shifted process \tilde{L} . The proof of Proposition 2.8 is provided in the Appendix. To verify H3 the following result can be applied.

Proposition 2.9 ([DS15, Theorem 3.1(c)]). *Assume that for some $\alpha \in (0, 2)$, $C > 0$ the symbol Φ satisfies*

$$|\Phi(\lambda)| \leq C|\lambda|^\alpha, \quad \lambda \in \mathbb{R}^d. \quad (2.10)$$

Then for any $p \in (0, \alpha)$, there exists $N = N(\kappa, \alpha, C, d)$ such that

$$\mathbb{E}|L_t|^p \leq Nt^{p/\alpha}, \quad t \in (0, 1].$$

One can also derive an explicit formula for the moments of L in terms of Φ [DS15, p. 3865]; this is also very useful in verifying H3. Namely, for any $p \in (0, 2)$, there exists a constant $N = N(p, d)$ such that for any $t > 0$

$$\mathbb{E}|L_t|^p = N \int_{\mathbb{R}^d} (1 - \operatorname{Re} e^{-t\Phi(\lambda)}) |\lambda|^{-p-d} d\lambda.$$

Now let us give an extensive list of examples of Lévy processes satisfying H1–H3. This list is inspired by [Pri12, p. 425], [Pri18, Example 6.2], [SSW12, p. 1147] and [CSZ18, Section 4]. All the corresponding proofs are placed in the Appendix.

Example 2.10 (General non-degenerate α -stable process, $\alpha \in (0, 2)$). Take $Q = 0$, $a = 0$, and

$$\nu(D) = \int_0^\infty \int_{\mathbb{S}} r^{-1-\alpha} \mathbb{1}_D(r\xi) \mu(d\xi) dr, \quad D \in \mathcal{B}(\mathbb{R}^d),$$

where μ is a finite non-negative measure concentrated on the unit sphere $\mathbb{S} := \{y \in \mathbb{R}^d : |y| = 1\}$ that is non-degenerate, i.e. its support is not contained in a proper linear subspace of \mathbb{R}^d . Then there exists $\kappa \in \mathbb{R}^d$ such that the shifted process $\tilde{L}_t := L_t + \kappa t$ satisfies H1–H3 (and, thus, Theorem 2.2 holds by Remark 2.6). If $\alpha \in [1, 2)$, then $\kappa = 0$; if $\alpha \in (0, 1)$, then $\kappa = \int_{|y| \leq 1} y \nu(dy) < \infty$

Example 2.11 (Standard isotropic d -dimensional α -stable process, $\alpha \in (0, 2)$). This is a special case of Example 2.10 with μ being the uniform measure on \mathbb{S} . One can check that in this case $\nu(D) = c_\alpha \int_D |y|^{-d-\alpha} dy$, $D \in \mathcal{B}(\mathbb{R}^d)$ and $\kappa = 0$ for all values of $\alpha \in (0, 2)$.

Example 2.12 (Cylindrical α -stable process, $\alpha \in (0, 2)$). Let L be a d -dimensional process whose coordinates are independent standard 1-dimensional α -stable processes. Then L satisfies H1–H3.

Example 2.13 (α -stable-type process, $\alpha \in (0, 2)$, see [SSW12, p. 1146]). Take $Q = 0$, $a = 0$, and

$$\nu(D) = \int_0^\infty \int_{\mathbb{S}} r^{-1-\alpha} \rho(r) \mathbb{1}_D(r\xi) \mu(d\xi) dr, \quad D \in \mathcal{B}(\mathbb{R}^d), \quad (2.11)$$

where μ is a symmetric (that is $\mu(D) = \mu(-D)$ for any Borel set D) non-degenerate finite non-negative measure concentrated on \mathbb{S} and $\rho: (0, \infty) \rightarrow \mathbb{R}_+$ is a measurable function such that for some constants $C, C_1, C_2 > 0$ one has

$$\mathbb{1}_{[0, C]}(r) \leq C_1 \rho(r) \leq C_2. \quad (2.12)$$

Then L satisfies H1–H3.

Example 2.14 (α -stable tempered process, $\alpha \in (0, 2)$). This is a special case of Example 2.13 with $\rho(r) = e^{-cr}$ with some $c > 0$.

Example 2.15 (Truncated α -stable process, $\alpha \in (0, 2)$). This is a special case of Example 2.13 with $\rho_r = c \mathbb{1}_{[0, 1]}$ with some $c > 0$ and μ being the uniform measure on \mathbb{S} . One can check that in this case $\nu(D) = c_\alpha \int_{D \cap \{y: |y| \leq 1\}} |y|^{-d-\alpha} dy$, $D \in \mathcal{B}(\mathbb{R}^d)$

Example 2.16 (Relativistic α -stable process, $\alpha \in (0, 2)$, [CMS90, CVST20]). Take L with symbol $\Phi(\lambda) = (|\lambda|^2 + C^{2/\alpha})^{\alpha/2} - C$, with some parameter $C > 0$. It satisfies H1–H3.

Example 2.17 (Brownian motion). If L is the standard d -dimensional Brownian motion, then it satisfies H1–H3 with $\alpha = 2$.

Example 2.18 (Linear combinations). (i) Let $L^{(1)}, L^{(2)}$ be two independent Lévy processes. Assume that the process $L^{(i)}$, $i = 1, 2$, satisfies H1–H3 with $\alpha = \alpha_i$. Then the process $L^{(1)} + L^{(2)}$ satisfies H1–H3 with $\alpha = \alpha_1 \vee \alpha_2$.

(ii) In particular, if L is a sum of a d -dimensional Brownian motion and the standard α -stable process, then the rate in (2.4) is $n^{-(\frac{1}{2} + \frac{\beta}{2} \wedge \frac{1}{2}) + \varepsilon}$ for $b \in \mathcal{C}^\beta$, $\beta > 0$.

- (iii) In general, if L has a non-degenerate diffusion part (that is, Q is positive definite) and for some $\gamma > 0$ we have $\int_{\{|y|>1\}} |y|^\gamma \nu(dy) < \infty$, then the statement of Theorem 2.2 is valid with $\alpha = 2$. This corresponds to the rate $n^{-(\frac{1}{2} + \frac{\beta}{2} \wedge \frac{1}{2}) + \varepsilon}$. This significantly improves [KS19b, Remark 2.3].

Remark 2.19. Let $d = 1$ and take $Q = 0$, $\nu(D) = c_\alpha \int_D y^{-1-\alpha} \mathbb{1}(y > 0) dy$, $D \in \mathcal{B}(\mathbb{R})$, $\alpha \in (0, 1)$, $a = \int_{|y| \leq 1} y \nu(dy) < \infty$. Then the process L satisfies H1–H3 (since it is a special case of (2.10)) and is increasing ([App09, Theorem 1.3.15]). Thus, regularization by noise can occur for monotone drivers as well.

Finally, let us present a simple example of a class of Lévy processes for which regularization by noise cannot occur. Suppose that $a = 0$, $Q = 0$, $\nu(\mathbb{R}^d) < \infty$. Then, the corresponding process L is a pure jump process and it has only finitely many jumps on the interval $[0, 1]$ [Sat13, Theorem 21.3]. Denoting its first jumping time by $T(\omega)$, we see that on the time interval $[0, T(\omega)]$, equation (1.1) becomes $dX_t = b(X_t)dt$. Clearly, if $b \in \mathcal{C}^\beta$, $\beta < 1$, this equation might have infinitely many or no solutions. Thus, for such Lévy noises the original equation (1.1) is not well-posed.

3. Preliminaries

Before we proceed to the proofs of our main results, let us collect a number of useful methods and bounds which we are going to apply later.

For a random variable ξ , a sub- σ -algebra $\mathcal{G} \subset \mathcal{F}$, and $p \geq 1$ we introduce the quantity

$$\|\xi\|_{L_p(\Omega)|\mathcal{G}} := (\mathbb{E}[|\xi|^p|\mathcal{G}])^{\frac{1}{p}}, \quad (3.1)$$

which is a \mathcal{G} -measurable non-negative random variable. It is clear that

$$\|\xi\|_{L_p(\Omega)} = \| \|\xi\|_{L_p(\Omega)|\mathcal{G}} \|_{L_p(\Omega)}.$$

Note that if $p \geq 1$, $\mathcal{G} \subset \mathcal{H}$ are σ -algebras, then the following simple bounds hold almost surely

$$\|\mathbb{E}[\xi|\mathcal{H}]\|_{L_p(\Omega)} \leq \|\xi\|_{L_p(\Omega)}; \quad \|\mathbb{E}[\xi|\mathcal{H}]\|_{L_p(\Omega)|\mathcal{G}} \leq \|\xi\|_{L_p(\Omega)|\mathcal{G}}. \quad (3.2)$$

These quantities are not norms, but rather \mathcal{G} -measurable nonnegative random variables. To simplify the presentation, any inequality between such expressions is understood in the almost sure sense.

3.1 Conditional shifted stochastic sewing lemma

An important tool to obtain Theorem 2.2 is an adjusted version of [Ger20, Lemma 2.2], which in turn is based on Lê's stochastic sewing lemma [Lê20]. We need the following notation. For $0 \leq S \leq T$ we denote a modified simplex

$$\Delta_{[S,T]} := \{(s, t) : S \leq s < t \leq T, s - (t - s) \geq S\}. \quad (3.3)$$

For a function $f : \Delta_{[S,T]} \rightarrow \mathbb{R}^d$ and a triplet of times (s, u, t) such that $S \leq s \leq u \leq t \leq T$, we denote

$$\delta f_{s,u,t} := f_{s,t} - f_{s,u} - f_{u,t}.$$

The conditional expectation given \mathcal{F}_s is denoted by \mathbb{E}^s .

Lemma 3.1. *Let $0 \leq S < T \leq 1$, $p \in [2, \infty)$ and let $(A_{s,t})_{(s,t) \in \Delta_{[S,T]}}$ be a family of random variables in $L_p(\Omega, \mathbb{R}^d)$ such that $A_{s,t}$ is \mathcal{F}_t -measurable. Let $\mathcal{G} \subset \mathcal{F}_S$ be a σ -algebra. Suppose that for some $\varepsilon_1, \varepsilon_2, \varepsilon_3 > 0$ and \mathcal{G} -measurable random variables $\Gamma_1, \Gamma_2, \Gamma_3 \geq 0$ the bounds*

$$\|A_{s,t}\|_{L_p(\Omega)|\mathcal{G}} \leq \Gamma_1 |t-s|^{1/2+\varepsilon_1}, \quad (3.4)$$

$$\|\mathbb{E}^{s-(t-s)} \delta A_{s,u,t}\|_{L_p(\Omega)|\mathcal{G}} \leq \Gamma_2 |t-s|^{1+\varepsilon_2} + \Gamma_3 |t-s|^{1+\varepsilon_3} \quad (3.5)$$

hold for all $(s,t) \in \Delta_{[S,T]}$ and $u = (s+t)/2$.

Further, suppose that there exists a process $\mathcal{A} = \{\mathcal{A}_t : t \in [S, T]\}$ such that for any $S \leq s \leq t \leq T$ one has

$$\mathcal{A}_t - \mathcal{A}_s = \lim_{m \rightarrow \infty} \sum_{i=1}^{m-1} A_{s+i\frac{t-s}{m}, s+(i+1)\frac{t-s}{m}} \text{ in probability.} \quad (3.6)$$

Then there exist deterministic constants $K_1, K_2 > 0$, which depend only on $\varepsilon_1, \varepsilon_2, \varepsilon_3, p$, and d such that for any $S \leq s \leq t \leq T$ we have

$$\|\mathcal{A}_t - \mathcal{A}_s\|_{L_p(\Omega)|\mathcal{G}} \leq K_1 \Gamma_1 |t-s|^{1/2+\varepsilon_1} + K_2 \Gamma_2 |t-s|^{1+\varepsilon_2} + K_3 \Gamma_3 |t-s|^{1+\varepsilon_3}. \quad (3.7)$$

The proof of Lemma 3.1 is given in the appendix.

3.2 Weighted John-Nirenberg inequality

Proposition 3.2. *Let $0 \leq S \leq T$ and let $\mathcal{A}: \Omega \times [S, T] \rightarrow \mathbb{R}^d$, $\xi: \Omega \times [S, T] \rightarrow \mathbb{R}_+$ be stochastic processes adapted to the filtration $(\mathcal{F}_t)_{t \in [S, T]}$. Assume additionally that \mathcal{A} is continuous and $\mathcal{A}_t \in L_1(\Omega)$ for all $t \in [S, T]$. Suppose that for any $S \leq s \leq t \leq T$ one has*

$$\mathbb{E}^s |\mathcal{A}_t - \mathcal{A}_s| \leq \xi_s \quad a.s. \quad (3.8)$$

Then for any $p \geq 1$ there exists a constant $N = N(p)$ such that for any $S \leq s \leq t \leq T$

$$\left\| \sup_{r \in [s, t]} |\mathcal{A}_r - \mathcal{A}_s| \right\|_{L_p(\Omega)|\mathcal{F}_s} \leq N \left\| \sup_{r \in [s, t]} \xi_r \right\|_{L_p(\Omega)|\mathcal{F}_s}. \quad (3.9)$$

The above proposition is very close to [Lê22a, Theorem 1.3], and the only difference is that the condition (3.8) is imposed there for all stopping times $s(\omega), t(\omega) \in [S, T]$. On the other hand, we only assume that (3.8) holds for deterministic $s, t \in [S, T]$; this will be crucial later in the proofs of Theorems 2.2, 2.4 and 2.5. For the case when ξ is a constant, [Lê22b, Proposition 2.2] shows that if (3.8) is satisfied for deterministic s, t , then it is also satisfied for s, t being stopping times. In general case, this seems to be not true. Therefore, for the proof of Proposition 3.2, we adapt the argument from [Lê22a, Theorem 1.3]. This is done in the appendix.

3.3 Heat kernel and related bounds

Recall that \mathcal{P} is the Markov transition semigroup associated to the process L . Quite often we use the following simple observation. For any measurable bounded function $f: \mathbb{R}^d \rightarrow \mathbb{R}^d$, $0 \leq s \leq t$, and any \mathcal{F}_s -measurable random vector ξ one has

$$\mathbb{E}^s f(L_t + \xi) = \mathcal{P}_{t-s} f(L_s + \xi). \quad (3.10)$$

We now formulate some consequences of H1-H3 in the form that they are actually used in the proofs.

Proposition 3.3. *Suppose that H1 and H2 hold. Then for every $\varepsilon > 0$, $\beta \geq 0$, ρ as below, $\mu \in (\varepsilon, 1 + \varepsilon]$, $\mu \geq (\beta - \rho)/\alpha$ there exists a constant $N = N(\alpha, \beta, \mu, \varepsilon, \rho)$ such that the following holds for any $f \in \mathcal{C}^\beta$:*

$$\|\mathcal{P}_t f\|_{\mathcal{C}^\rho} \leq N \|f\|_{\mathcal{C}^\beta} t^{\frac{(\beta-\rho)\wedge 0}{\alpha}}, \quad 0 < t \leq 1, \quad \rho \geq 0; \quad (3.11)$$

$$\|\mathcal{P}_t f - \mathcal{P}_s f\|_{\mathcal{C}^\rho} \leq N \|f\|_{\mathcal{C}^\beta} s^{\frac{\beta-\rho}{\alpha}-\mu} (t-s)^{\mu-\varepsilon}, \quad 0 \leq s < t \leq 1, \quad \rho = 0, 1. \quad (3.12)$$

The proof of this proposition is mostly technical and is provided in the appendix.

Assuming that L satisfies H3, it is immediate to see that this implies that for any $m > 0$, $p \geq 1$, $\varepsilon > 0$ there exists a constant $N = N(p, m, \varepsilon, \alpha)$ such that

$$\| |L_t|^m \wedge 1 \|_{L_p(\Omega)} \leq N t^{(\frac{m}{\alpha} \wedge \frac{1}{p}) - \varepsilon}, \quad t \in [0, 1]. \quad (3.13)$$

We further recall two elementary inequalities:

$$|f(x_1) - f(x_2)| \leq |x_1 - x_2| \|f\|_{\mathcal{C}^1} \quad (3.14)$$

$$|f(x_1) - f(x_2) - f(x_3) + f(x_4)| \leq |x_1 - x_2 - x_3 + x_4| \|f\|_{\mathcal{C}^1} + |x_1 - x_2| |x_1 - x_3| \|f\|_{\mathcal{C}^2} \quad (3.15)$$

for any $x_1, x_2, x_3, x_4 \in \mathbb{R}^d$ and any f from \mathcal{C}^1 or \mathcal{C}^2 , respectively.

4. Proof of the main results

We will denote

$$\varphi := X - L, \quad \varphi^n = X^n - L, \quad n \in \mathbb{Z}_+. \quad (4.1)$$

We consider the following decomposition of the difference between the additive functional of the process and its estimate. For $0 \leq s \leq t \leq 1$, $n \in \mathbb{Z}_+$, $f \in \mathcal{C}^\beta$ we write

$$\begin{aligned} \int_s^t f(X_r) dr - \int_s^t f(X_{\kappa_n(r)}^n) dr &= \int_s^t f(L_r + \varphi_r) - f(L_r + \varphi_r^n) dr \\ &\quad + \int_s^t f(L_r + \varphi_r^n) - f(L_r + \varphi_{\kappa_n(r)}^n) dr \\ &\quad + \int_s^t f(L_r + \varphi_{\kappa_n(r)}^n) - f(L_{\kappa_n(r)} + \varphi_{\kappa_n(r)}^n) dr \\ &=: \mathcal{E}_{s,t}^{f,n,1} + \mathcal{E}_{s,t}^{f,n,2} + \mathcal{E}_{s,t}^{f,n,3}. \end{aligned} \quad (4.2)$$

Clearly, for $f = b$ we get the increment of the difference between the process and its Euler approximation:

$$(X_t - X_t^n) - (X_s - X_s^n) = \int_s^t b(X_r) dr - \int_s^t b(X_{\kappa_n(r)}^n) dr = \mathcal{E}_{s,t}^{b,n,1} + \mathcal{E}_{s,t}^{b,n,2} + \mathcal{E}_{s,t}^{b,n,3}. \quad (4.3)$$

Remark 4.1. This decomposition differs from the one in e.g. [BDG21a]: therein, $\mathcal{E}^{b,n,2} + \mathcal{E}^{b,n,3}$ can be treated as one term, and by Girsanov's theorem, the perturbation φ^n can in fact be transformed away. Such trick is not available in the Lévy case due to the lack of an appropriate Girsanov's theorem.

Our goal is to bound the $L_p(\Omega)$ norm of the left-hand side of (4.3). However, doing this directly for the term $\mathcal{E}_{s,t}^{b,n,3}$ would lead to a rate that deteriorates for large p , see Remark 4.8 below. Therefore, we bound a conditional L^2 -“norms” instead and eventually after buckling apply John-Nirenberg inequality. First, in Section 4.1 we prove certain a priori bounds for φ, φ^n . Then in Section 4.2 we produce general bounds for conditional “norms” of integrals of irregular functions along the perturbed Lévy process, which are then applied to bound $\mathcal{E}_{s,t}^{f,n,i}$ in Section 4.3. Finally, these bounds are combined for the proofs of the main theorems in Section 4.4.

4.1 A priori bounds

Let $f: [0, 1] \times \Omega \rightarrow \mathbb{R}^d$ be a measurable bounded function adapted to the filtration \mathbb{F} . Let $\gamma \in (0, 1]$, $p \geq 2$, $0 \leq S \leq T \leq 1$, let $\mathcal{G} \subset \mathcal{F}_S$ be a σ -algebra. We consider the following quantities of f :

$$\begin{aligned} \|f\|_{\mathcal{C}_p^0|\mathcal{G},[S,T]} &:= \sup_{t \in [S,T]} \|f(t)\|_{L_p(\Omega)|\mathcal{G}}; \\ [f]_{\mathcal{C}_p^\gamma|\mathcal{G},[S,T]} &:= \sup_{s,t \in [S,T]} \frac{\|f(t) - f(s)\|_{L_p(\Omega)|\mathcal{G}}}{|t - s|^\gamma}; \end{aligned} \quad (4.4)$$

$$[f]_{\mathbf{C}_p^\gamma|\mathcal{G},[S,T]} := \sup_{s,t \in [S,T]} \frac{\|\|f(t) - \mathbb{E}^s f(t)\|_{L_1(\Omega)|\mathcal{F}_s}\|_{L_p(\Omega)|\mathcal{G}}}{|t - s|^\gamma}. \quad (4.5)$$

Part (iii) of the following lemma is the main a priori estimate on the “stochastic regularity” of φ, φ^n .

Lemma 4.2. (i) *Let $q \geq 1$. Let $\mathcal{G} \subset \mathcal{F}$ be a σ -algebra. Let $Y, Z \in L_q(\Omega)$ be random variables and suppose that Z is \mathcal{G} -measurable. Then*

$$\|Y - \mathbb{E}[Y|\mathcal{G}]\|_{L_q(\Omega)|\mathcal{G}} \leq 2\|Y - Z\|_{L_q(\Omega)|\mathcal{G}}, \quad a.s. \quad (4.6)$$

(ii) *For any $0 \leq s \leq t \leq 1$, $q \geq 1$, measurable function $f: [s, t] \times \Omega \rightarrow \mathbb{R}^d$ adapted to the filtration \mathbb{F} , $\gamma \in (0, 1)$, σ -algebra $\mathcal{G} \subset \mathcal{F}_s$ one has*

$$[f]_{\mathbf{C}_q^\gamma|\mathcal{G},[s,t]} \leq 2[f]_{\mathcal{C}_q^\gamma|\mathcal{G},[s,t]}. \quad (4.7)$$

(iii) *Let $q \geq 1$, $\varepsilon > 0$. Assume that $\beta > 1 - \alpha$ and that H3 holds. There exists a constant $N = N(\beta, \|b\|_{C^\beta}, \alpha, \varepsilon, q, M)$ such that for $0 \leq s \leq t \leq 1$ one has a.s.*

$$\|\varphi_t - \mathbb{E}^s \varphi_t\|_{L_q(\Omega)|\mathcal{F}_s} \leq N|t - s|^{1 + (\frac{\beta \wedge 1}{\alpha} \wedge \frac{1}{q}) - \varepsilon}; \quad (4.8)$$

$$\|\varphi_t^n - \mathbb{E}^s \varphi_t^n\|_{L_q(\Omega)|\mathcal{F}_s} \leq N|t - s|^{1 + (\frac{\beta \wedge 1}{\alpha} \wedge \frac{1}{q}) - \varepsilon}. \quad (4.9)$$

Proof. (i) We have

$$\begin{aligned} \|Y - \mathbb{E}[Y|\mathcal{G}]\|_{L_q(\Omega)|\mathcal{G}} &\leq \|Y - Z\|_{L_q(\Omega)|\mathcal{G}} + \|\mathbb{E}[Y|\mathcal{G}] - Z\|_{L_q(\Omega)|\mathcal{G}} \\ &= \|Y - Z\|_{L_q(\Omega)|\mathcal{G}} + \|\mathbb{E}[Y - Z|\mathcal{G}]\|_{L_q(\Omega)|\mathcal{G}} \\ &\leq 2\|Y - Z\|_{L_q(\Omega)|\mathcal{G}}, \end{aligned}$$

where the last inequality follows from (3.2).

(ii) By part (i) of the lemma and Jensen's inequality we have for any $s \leq s' \leq t' \leq t$

$$\|f(t') - \mathbb{E}^{s'} f(t')\|_{L_1(\Omega)|\mathcal{F}_{s'}} \|_{L_q(\Omega)|\mathcal{G}} \leq 2 \|f(t') - f(s')\|_{L_1(\Omega)|\mathcal{F}_{s'}} \|_{L_q(\Omega)|\mathcal{G}} \leq 2 \|f(t') - f(s')\|_{L_q(\Omega)|\mathcal{G}},$$

where we used that $\mathcal{G} \subset \mathcal{F}_s \subset \mathcal{F}_{s'}$. The desired result follows now from the definitions of the seminorms (4.4) and (4.5).

(iii) Without loss of generality, we can assume that $\beta \leq 1$. Suppose that (4.8) holds for some $m \geq 0$ in place of $1 + (\frac{\beta \wedge 1}{\alpha} \wedge \frac{1}{q}) - \varepsilon$. This is certainly true for $m = 0$ thanks to the fact that b is bounded; we proceed now by induction on m . We apply (4.6) with $\mathcal{G} = \mathcal{F}_s$, $Y = \varphi_t$, $Z = \varphi_s + \int_s^t b(L_s + \mathbb{E}^s \varphi_r) dr$. We get

$$\begin{aligned} \|\varphi_t - \mathbb{E}^s \varphi_t\|_{L_q(\Omega)|\mathcal{F}_s} &\leq 2 \left\| \varphi_t - \varphi_s - \int_s^t b(L_s + \mathbb{E}^s \varphi_r) dr \right\|_{L_q(\Omega)|\mathcal{F}_s} \\ &= 2 \left\| \int_s^t (b(L_r + \varphi_r) - b(L_s + \mathbb{E}^s \varphi_r)) dr \right\|_{L_q(\Omega)|\mathcal{F}_s} \\ &\leq N \left\| \int_s^t (|L_r - L_s|^\beta + |\varphi_r - \mathbb{E}^s \varphi_r|^\beta) \wedge 1 dr \right\|_{L_q(\Omega)|\mathcal{F}_s} \\ &\leq N \int_s^t (\|L_r - L_s\|_{L_q(\Omega)|\mathcal{F}_s}^\beta + \|\varphi_r - \mathbb{E}^s \varphi_r\|_{L_q(\Omega)|\mathcal{F}_s}^\beta) dr. \end{aligned}$$

Using (3.13), the independence of $L_r - L_s$ from \mathcal{F}_s , and the induction hypothesis, we get a.s.

$$\|\varphi_t - \mathbb{E}^s \varphi_t\|_{L_q(\Omega)|\mathcal{F}_s} \leq N |t - s|^{1 + (\frac{\beta}{\alpha} \wedge \frac{1}{q} \wedge \beta m) - \varepsilon}.$$

It is elementary to see that if $\varepsilon > 0$ is small enough, then the recursion $m_0 = 0$, $m_{i+1} = 1 + (\frac{\beta}{\alpha} \wedge \frac{1}{q} \wedge \beta m_i) - \varepsilon$ reaches $1 + (\frac{\beta}{\alpha} \wedge \frac{1}{q}) - \varepsilon$ in finitely many steps (recall that $\alpha > 1 - \beta$ and thus $\alpha > (1 - \beta)/(1 - \varepsilon)$ for small enough $\varepsilon > 0$). Recalling our initial assumption $\beta \leq 1$, we get (4.8).

Inequality (4.9) is obtained by a similar argument, though one has to be a bit more careful because now $L_{\kappa_n(r)} - L_{\kappa_n(s)}$ is not independent of \mathcal{F}_s . For fixed $s \in [0, 1]$, define s' to be the smallest grid point which is bigger or equal to s , that is, $s' := \lceil ns \rceil n^{-1}$. It is crucial to note that $\varphi_{s'}^n$ is \mathcal{F}_s measurable.

We proceed by induction as before and assume that (4.9) holds for some $m \geq 0$. If $s \leq t < s'$, then φ_t^n is \mathcal{F}_s -measurable. Hence $\varphi_t^n = \mathbb{E}^s \varphi_t^n$ and the left-hand side of (4.9) is zero. Therefore it remains to consider the case $t \geq s'$. In this case, using again (4.6) with $\mathcal{G} = \mathcal{F}_s$, $Y = \varphi_t^n$, $Z = \varphi_{s'}^n + \int_{s'}^t b(L_s + \mathbb{E}^s \varphi_{\kappa_n(r)}^n) dr$, we deduce

$$\begin{aligned} \|\varphi_t^n - \mathbb{E}^s \varphi_t^n\|_{L_q(\Omega)|\mathcal{F}_s} &\leq 2 \left\| \varphi_t^n - \varphi_{s'}^n - \int_{s'}^t b(L_s + \mathbb{E}^s \varphi_{\kappa_n(r)}^n) dr \right\|_{L_q(\Omega)|\mathcal{F}_s} \\ &= 2 \left\| \int_{s'}^t (b(L_{\kappa_n(r)} + \varphi_{\kappa_n(r)}^n) - b(L_s + \mathbb{E}^s \varphi_{\kappa_n(r)}^n)) dr \right\|_{L_q(\Omega)|\mathcal{F}_s} \\ &\leq N \left\| \int_{s'}^t (|L_{\kappa_n(r)} - L_s|^\beta + |\varphi_{\kappa_n(r)}^n - \mathbb{E}^s \varphi_{\kappa_n(r)}^n|^\beta) \wedge 1 dr \right\|_{L_q(\Omega)|\mathcal{F}_s}. \end{aligned}$$

Note that for $r \geq s'$, $\kappa_n(r) \geq s' \geq s$, and therefore $L_{\kappa_n(r)} - L_s$ is independent of \mathcal{F}_s . From here we obtain (4.9) exactly as before. \square

Remark 4.3. The reason for the non-standard portion of our main assumption (2.3) (condition $\beta > 2 - 2\alpha$) and the strange threshold $2/3$ in Remark 2.7 is the appearance of $1/q$ in (4.8)-(4.9).

4.2 General bounds

Since $\mathcal{E}^{f,n,1}$ and $\mathcal{E}^{f,n,2}$ in decomposition (4.2) have similar forms (difference of averages of f along L with two different perturbations), we begin with the following general bound that can be applied to both.

Lemma 4.4. *Let $p \in [2, \infty)$. Assume H1-H3. Let $\tau, \gamma, \varepsilon_0 \in (0, 1]$, $\theta > 0$, be constants satisfying*

$$\theta \in (1 - \frac{\alpha}{2}, 2], \quad \frac{\theta - 2}{\alpha} + \tau > 0, \quad \gamma + \frac{\theta - 1}{\alpha} > 0. \quad (4.10)$$

Let $f \in \mathcal{C}^\theta$. Let $g, h: [0, 1] \times \Omega \rightarrow \mathbb{R}^d$ be bounded, adapted, measurable functions and suppose that there exist $C_g, C_h > 0$ such that for any $0 \leq s \leq t \leq 1$ one has a.s.

$$\mathbb{E}^s |g_t - \mathbb{E}^s g_t| \leq C_g |t - s|^\tau, \quad (4.11)$$

$$\|h_t - \mathbb{E}^s h_t\|_{L_1(\Omega)} \leq C_h |t - s|^{\varepsilon_0}. \quad (4.12)$$

Then there exists a constant $N = N(\alpha, \theta, p, d, \gamma, \tau)$ such that for any $0 \leq S \leq T \leq 1$ and any σ -algebra $\mathcal{G} \subset \mathcal{F}_S$ one has the bound

$$\begin{aligned} \left\| \int_S^T f(L_r + g_r) - f(L_r + h_r) dr \right\|_{L_p(\Omega)|\mathcal{G}} &\leq N \|f\|_{\mathcal{C}^\theta} (T - S)^{1 + \frac{(\theta-1) \wedge 0}{\alpha}} \|g - h\|_{\mathcal{C}_p^0|\mathcal{G}, [S, T]} \\ &\quad + N \|f\|_{\mathcal{C}^\theta} (T - S)^{\gamma + \frac{(\theta-1) \wedge 0}{\alpha} + 1} [g - h]_{\mathcal{C}_p^\gamma|\mathcal{G}, [S, T]} \\ &\quad + N \|f\|_{\mathcal{C}^\theta} C_g (T - S)^{1 + \frac{\theta-2}{\alpha} + \tau} \|g - h\|_{\mathcal{C}_p^0|\mathcal{G}, [S, T]}. \end{aligned} \quad (4.13)$$

Remark 4.5. It is pivotal that the seminorm appearing in the right-hand side of (4.13) is $[g - h]_{\mathcal{C}_p^\gamma|\mathcal{G}, [S, T]}$ rather than a much less precise seminorm $[g - h]_{\mathcal{C}_p^\gamma|\mathcal{G}, [S, T]}$ (recall Lemma 4.2(ii) and the definitions of the seminorms in (4.5) and (4.4)). This will be crucial for bounding $\mathcal{E}^{f,n,2}$, see Remark 4.10 below.

Proof. Fix $0 \leq S \leq T \leq 1$. Put

$$\begin{aligned} A_{s,t} &:= \mathbb{E}^{s-(t-s)} \int_s^t f(L_r + \mathbb{E}^{s-(t-s)} g_r) - f(L_r + \mathbb{E}^{s-(t-s)} h_r) dr, \quad (s, t) \in \Delta_{[S, T]}; \\ \mathcal{A}_t &:= \int_0^t f(L_r + g_r) - f(L_r + h_r) dr, \quad t \in [S, T]. \end{aligned}$$

Let us verify that the processes A, \mathcal{A} satisfy all the conditions of the stochastic sewing lemma (Lemma 3.1).

Let $(s, t) \in \Delta_{[S, T]}$. Then recalling (3.10) and (3.14), we see that

$$\begin{aligned} |A_{s,t}| &\leq \int_s^t |\mathcal{P}_{r-(s-(t-s))} f(L_{s-(t-s)} + \mathbb{E}^{s-(t-s)} g_r) - \mathcal{P}_{r-(s-(t-s))} f(L_{s-(t-s)} + \mathbb{E}^{s-(t-s)} h_r)| dr \\ &\leq \int_s^t \|\mathcal{P}_{r-(s-(t-s))} f\|_{\mathcal{C}^1} |\mathbb{E}^{s-(t-s)}(g_r - h_r)| dr. \end{aligned}$$

Thus, by (3.11) (applied with $\rho = 1$ and $\beta = \theta$) and (3.2), we have

$$\|A_{s,t}\|_{L_p(\Omega)|\mathcal{G}} \leq \|f\|_{\mathcal{C}^\theta} \int_s^t (r - s)^{\frac{(\theta-1) \wedge 0}{\alpha}} \|\mathbb{E}^{s-(t-s)}(g_r - h_r)\|_{L_p(\Omega)|\mathcal{G}} dr$$

$$\begin{aligned}
&\leq N\|f\|_{\mathcal{C}^\theta}(t-s)^{1+\frac{(\theta-1)\wedge 0}{\alpha}} \sup_{r \in [S, T]} \|\mathbb{E}^{s-(t-s)}(g_r - h_r)\|_{L_p(\Omega)|\mathcal{G}} \\
&\leq N\|f\|_{\mathcal{C}^\theta}(t-s)^{1+\frac{(\theta-1)\wedge 0}{\alpha}} \sup_{r \in [S, T]} \|g_r - h_r\|_{L_p(\Omega)|\mathcal{G}} \\
&= N\|f\|_{\mathcal{C}^\theta}(t-s)^{1+\frac{(\theta-1)\wedge 0}{\alpha}} \|g - h\|_{\mathcal{C}_p^0|\mathcal{G}, [S, T]}.
\end{aligned}$$

Here in the penultimate line we used that $\mathcal{G} \subset \mathcal{F}_S \subset \mathcal{F}_{s-(t-s)}$. Note that by the assumption $\theta > 1 - \alpha/2$, we have $1 + \frac{\theta-1}{\alpha} > 1/2$. Therefore, condition (3.4) is satisfied with $\Gamma_1 = N\|f\|_{\mathcal{C}^\theta} \|g - h\|_{\mathcal{C}_p^0|\mathcal{G}, [S, T]}$.

Now let us verify condition (3.5). As required, we take $(s, t) \in \Delta_{[S, T]}$, and $u := (t+s)/2$. It will be convenient to denote $s_1 := s - (t-s)$, $s_2 := s - (u-s)$, $s_3 := s$, $s_4 := u$, $s_5 := t$. One has $s_1 \leq s_2 \leq s_3 \leq s_4 \leq s_5$. Then we deduce

$$\begin{aligned}
&\mathbb{E}^{s-(t-s)} \delta A_{s, u, t} \\
&= \mathbb{E}^{s_1} \delta A_{s_3, s_4, s_5} \\
&= \mathbb{E}^{s_1} \int_{s_3}^{s_4} (f(L_r + \mathbb{E}^{s_1} g_r) - f(L_r + \mathbb{E}^{s_1} h_r) - f(L_r + \mathbb{E}^{s_2} g_r) - f(L_r + \mathbb{E}^{s_2} h_r)) dr \\
&\quad + \mathbb{E}^{s_1} \int_{s_4}^{s_5} (f(L_r + \mathbb{E}^{s_1} g_r) - f(L_r + \mathbb{E}^{s_1} h_r) - f(L_r + \mathbb{E}^{s_3} g_r) - f(L_r + \mathbb{E}^{s_3} h_r)) dr \\
&=: I_1 + I_2.
\end{aligned} \tag{4.14}$$

Here in the term I_2 we used the identity $u - (t-s) = s = s_3$. We begin with the analysis of I_1 . Recalling (3.10), we obviously have

$$\begin{aligned}
I_1 &= \mathbb{E}^{s_1} \mathbb{E}^{s_2} \int_{s_3}^{s_4} (f(L_r + \mathbb{E}^{s_1} g_r) - f(L_r + \mathbb{E}^{s_1} h_r) - f(L_r + \mathbb{E}^{s_2} g_r) - f(L_r + \mathbb{E}^{s_2} h_r)) dr \\
&= \mathbb{E}^{s_1} \int_{s_3}^{s_4} (\mathcal{P}_{r-s_2} f(L_{s_2} + \mathbb{E}^{s_1} g_r) - \mathcal{P}_{r-s_2} f(L_{s_2} + \mathbb{E}^{s_1} h_r) \\
&\quad - \mathcal{P}_{r-s_2} f(L_{s_2} + \mathbb{E}^{s_2} g_r) - \mathcal{P}_{r-s_2} f(L_{s_2} + \mathbb{E}^{s_2} h_r)) dr.
\end{aligned}$$

Applying (3.15) and (3.11) we see that

$$\begin{aligned}
|I_1| &\leq \|f\|_{\mathcal{C}^\theta} \int_{s_3}^{s_4} (r - s_2)^{\frac{(\theta-1)\wedge 0}{\alpha}} \mathbb{E}^{s_1} |\mathbb{E}^{s_1}(g_r - h_r) - \mathbb{E}^{s_2}(g_r - h_r)| dr \\
&\quad + \|f\|_{\mathcal{C}^\theta} \int_{s_3}^{s_4} (r - s_2)^{\frac{\theta-2}{\alpha}} |\mathbb{E}^{s_1}(g_r - h_r)| \mathbb{E}^{s_1} |\mathbb{E}^{s_2} g_r - \mathbb{E}^{s_1} g_r| dr.
\end{aligned} \tag{4.15}$$

Using conditional Jensen's inequality and the assumption (4.11), we see that a.s.

$$\mathbb{E}^{s_1} |\mathbb{E}^{s_2} g_r - \mathbb{E}^{s_1} g_r| = \mathbb{E}^{s_1} |\mathbb{E}^{s_2}(g_r - \mathbb{E}^{s_1} g_r)| \leq \mathbb{E}^{s_1} |g_r - \mathbb{E}^{s_1} g_r| \leq C_g |r - s_1|^\tau. \tag{4.16}$$

Similarly,

$$\begin{aligned}
\mathbb{E}^{s_1} |\mathbb{E}^{s_2}(g_r - h_r) - \mathbb{E}^{s_1}(g_r - h_r)| &= \mathbb{E}^{s_1} |\mathbb{E}^{s_2}((g_r - h_r) - \mathbb{E}^{s_1}(g_r - h_r))| \\
&\leq \mathbb{E}^{s_1} |(g_r - h_r) - \mathbb{E}^{s_1}(g_r - h_r)|.
\end{aligned} \tag{4.17}$$

Combining (4.15), (4.16), (4.17), and using the Minkowski inequality together with (3.2) and the fact that $\mathcal{G} \subset \mathcal{F}_{s_1}$, we finally get

$$\|I_1\|_{L_p(\Omega)|\mathcal{G}} \leq N\|f\|_{\mathcal{C}^\theta}(s_4 - s_1)^{\gamma + \frac{(\theta-1)\wedge 0}{\alpha} + 1} [g - h]_{\mathbf{C}_p^\gamma|\mathcal{G}, [S, T]}$$

$$\begin{aligned}
& + N\|f\|_{\mathcal{C}^\theta} C_g(s_3 - s_2)^{\frac{\theta-2}{\alpha}} (s_4 - s_3)(s_4 - s_1)^\tau \|g - h\|_{\mathcal{C}_p^0[\mathcal{G}, [S, T]]} \\
& \leq N\|f\|_{\mathcal{C}^\theta} (t - s)^{\gamma + \frac{(\theta-1)\wedge 0}{\alpha} + 1} [g - h]_{\mathcal{C}_p^\gamma[\mathcal{G}, [S, T]]} \\
& + N\|f\|_{\mathcal{C}^\theta} C_g(t - s)^{\frac{\theta-2}{\alpha} + 1 + \tau} \|g - h\|_{\mathcal{C}_p^0[\mathcal{G}, [S, T]]}, \tag{4.18}
\end{aligned}$$

where the last inequality follows from the fact $s_3 - s_2 = s_4 - s_3 = (t - s)/2$ and $s_4 - s_1 = (u - s) + (t - s) = \frac{3}{2}(t - s)$. By exactly the same argument (we just need to take s_3 in place of s_2 , s_4 in place of s_3 and s_5 in place of s_4), we get the exact same bound for I_2 , and then by (4.14), for $\mathbb{E}^{s-(t-s)}\delta A_{s,u,t}$ as well. Since by the assumptions of the lemma $\gamma + \frac{(\theta-1)\wedge 0}{\alpha} + 1 > 1$ and $\frac{\theta-2}{\alpha} + 1 + \tau > 1$, we see that condition (3.5) is satisfied with $\Gamma_2 = N\|f\|_{\mathcal{C}^\theta} [g - h]_{\mathcal{C}_p^\gamma[\mathcal{G}, [S, T]]}$ and $\Gamma_3 = N\|f\|_{\mathcal{C}^\theta} C_g \|g - h\|_{\mathcal{C}_p^0[\mathcal{G}, [S, T]]}$.

It remains to verify condition (3.6). Let $s, t \in [S, T]$, $s < t$. Fix $m \in \mathbb{N}$. Denote $t_i := s + i\frac{t-s}{m}$, $i = 0, \dots, m$. Note that $t_i - (t_{i+1} - t_i) = t_{i-1}$. Then we have

$$\begin{aligned}
|\mathcal{A}_t - \mathcal{A}_s - \sum_{i=1}^{m-1} A_{t_i, t_{i+1}}| & \leq \sum_{i=1}^{m-1} \int_{t_i}^{t_{i+1}} |f(L_r + g_r) - \mathcal{P}_{r-t_{i-1}} f(L_{t_{i-1}} + \mathbb{E}^{t_{i-1}} g_r)| dr \\
& + \sum_{i=1}^{m-1} \int_{t_i}^{t_{i+1}} |f(L_r + h_r) - \mathcal{P}_{r-t_{i-1}} f(L_{t_{i-1}} + \mathbb{E}^{t_{i-1}} h_r)| dr \\
& + \int_{t_0}^{t_1} (|f(L_r + g_r)| + |f(L_r + h_r)|) dr \\
& =: I_{m,1} + I_{m,2} + I_{m,3}. \tag{4.19}
\end{aligned}$$

Using (3.12) (with $\rho = 0$, $\theta \wedge \alpha$ in place of β , and $\mu = \frac{\theta \wedge \alpha}{\alpha}$) we easily deduce that for any $\varepsilon > 0$

$$\begin{aligned}
& |f(L_r + g_r) - \mathcal{P}_{r-t_{i-1}} f(L_{t_{i-1}} + \mathbb{E}^{t_{i-1}} g_r)| \\
& \leq \|f\|_{\mathcal{C}^{\theta \wedge 1}} (|L_r - L_{t_{i-1}}|^{\theta \wedge 1} \wedge 1) + \|f\|_{\mathcal{C}^{\theta \wedge 1}} |g_r - \mathbb{E}^{t_{i-1}} g_r|^{\theta \wedge 1} + \|f\|_{\mathcal{C}^{\theta \wedge \alpha}} |r - t_{i-1}|^{\frac{\theta \wedge \alpha}{\alpha} - \varepsilon}.
\end{aligned}$$

This together with (3.13), (4.11) and the Minkowski inequality yields

$$\|I_{m,1}\|_{L_1(\Omega)} \leq N(1 + C_g) \|f\|_{\mathcal{C}^\theta} (t - s) \frac{1}{m^{(\frac{\theta \wedge 1}{\alpha} \wedge ((\theta \wedge 1)\tau) \wedge 1) - \varepsilon}}. \tag{4.20}$$

Similarly, with the help of (4.12) we bound

$$\|I_{m,2}\|_{L_1(\Omega)} \leq N(1 + C_h) \|f\|_{\mathcal{C}^\theta} (t - s) \frac{1}{m^{(\frac{\theta \wedge 1}{\alpha} \wedge ((\theta \wedge 1)\varepsilon_0) \wedge 1) - \varepsilon}}. \tag{4.21}$$

Finally, it is obvious that $|I_{m,3}| \leq Nm^{-1} \|f\|_{\mathcal{C}^0}$. Therefore, substituting this, (4.20) and (4.21) into (4.19) and choosing $\varepsilon > 0$ sufficiently small, we see that the sum $\sum_{i=1}^{m-1} A_{t_i, t_{i+1}}$ converges to $\mathcal{A}_t - \mathcal{A}_s$ in $L_1(\Omega)$ and hence in probability as $m \rightarrow \infty$. Hence, (3.6) holds.

Thus, all the conditions of Lemma 3.1 are satisfied. The claimed bound (4.13) follows now from (3.7). \square

Remark 4.6. We now understand why it was essential to use the shifted stochastic sewing lemma rather than the usual stochastic sewing lemma. Indeed, the exponent in the second

term in (4.15), $\frac{\theta-2}{\alpha}$, can be less than -1 . Had we applied the usual stochastic sewing lemma, we would have been required to impose $\frac{\theta-2}{\alpha} > -1$ to ensure integrability. Later, when we apply Lemma 4.4 for $f = b$ and $\theta = \beta$, this would have led to the suboptimal condition $\beta > 2 - \alpha$ rather than $\beta > 1 - \frac{\alpha}{2}$. This issue is effectively resolved through the use of shifting.

Next, we obtain a general quadrature estimate. It will be crucial for bounding the third error term $\mathcal{E}^{f,n,3}$ in decomposition (4.2) as well as for the proof of Theorem 2.4. An analogue of such a bound in the case of fractional Brownian motion in place of L and 0 in place of g is obtained in [BDG21a, Lemma 4.1], with rate that is consistent with (4.25) below.

Lemma 4.7. *Assume H1-H3. Let $g: [0, 1] \times \Omega \rightarrow \mathbb{R}^d$ be a bounded measurable function, let $f \in \mathcal{C}^\theta$, $\theta \in [0, 1]$. Suppose that the following holds:*

(i) *there exist constants $\tau > 0$, $C_g > 0$ such that*

$$\tau > \frac{1}{2} + \frac{1}{\alpha} + \frac{\theta}{\alpha} \wedge \frac{1}{2} - \frac{\theta}{\alpha}; \quad (4.22)$$

$$\mathbb{E}^s |\mathbb{E}^t g_r - \mathbb{E}^s g_r| \leq C_g |r - s|^\tau, \quad 0 \leq s \leq r \leq 1, \quad t \in [s, 1]; \quad (4.23)$$

(ii) *for some $n \in \mathbb{N}$ and all $t \in [0, 1]$*

$$g_{\kappa_n(t)} \text{ is } \mathcal{F}_{(\kappa_n(t) - \frac{1}{n}) \vee 0} \text{ measurable.} \quad (4.24)$$

(iii) *$g \equiv 0$ or $\theta > 0$.*

Then for any $\varepsilon \in (0, 1/2)$ there exists a constant $N = N(\alpha, \theta, d, \tau, \varepsilon, M)$ independent of n such that for any $\{0 \leq S \leq T \leq 1\}$, and any σ -algebra $\mathcal{G} \subset \mathcal{F}_{\kappa_n(S)}$ the following holds:

$$\begin{aligned} & \left\| \int_S^T (f(L_r + g_{\kappa_n(r)}) - f(L_{\kappa_n(r)} + g_{\kappa_n(r)})) dr \right\|_{L_2(\Omega)|\mathcal{G}} \\ & \leq N \|f\|_{\mathcal{C}^\theta} (1 + C_g) n^{-(\frac{1}{2} + \frac{\theta}{\alpha} \wedge \frac{1}{2}) + 2\varepsilon} |T - S|^{\frac{1}{2} + \varepsilon}. \end{aligned} \quad (4.25)$$

Proof. To simplify notations, set $\psi_s := g_{\kappa_n(s)}$. We fix $0 \leq S \leq T \leq 1$ and apply Lemma 3.1 for the processes

$$A_{s,t} := \mathbb{E}^{s-(t-s)} \int_s^t (f(L_r + \mathbb{E}^{s-(t-s)} \psi_r) - f(L_{\kappa_n(r)} + \mathbb{E}^{s-(t-s)} \psi_r)) dr, \quad (s, t) \in \Delta_{[S,T]}; \quad (4.26)$$

$$\mathcal{A}_t := \int_0^t (f(L_r + \psi_r) - f(L_{\kappa_n(r)} + \psi_r)) dr, \quad t \in [S, T].$$

First we verify (3.4). If $(s, t) \in \Delta_{[S,T]}$ and $s \leq t \leq s + 2/n$, then we have from (3.13) (with $m = \theta$ and $p = 2$)

$$\|A_{s,t}\|_{L_2(\Omega)|\mathcal{G}} \leq \int_s^t \|f(L_r + \mathbb{E}^{s-(t-s)} \psi_r) - f(L_{\kappa_n(r)} + \mathbb{E}^{s-(t-s)} \psi_r)\|_{L_2(\Omega)|\mathcal{G}} dr$$

$$\begin{aligned}
&\leq \|f\|_{\mathcal{C}^\theta} \int_s^t \| |L_r - L_{\kappa_n(r)}|^\theta \wedge 1 \|_{L_2(\Omega)|\mathcal{G}} dr \\
&\leq N \|f\|_{\mathcal{C}^\theta} |t-s| n^{-(\frac{\theta}{\alpha} \wedge \frac{1}{2}) + \varepsilon} \\
&\leq N \|f\|_{\mathcal{C}^\theta} |t-s|^{\frac{1}{2} + \varepsilon} n^{-(\frac{1}{2} + \frac{\theta}{\alpha} \wedge \frac{1}{2}) + 2\varepsilon},
\end{aligned} \tag{4.27}$$

where we used in the penultimate inequality that $\mathcal{G} \subset \mathcal{F}_{\kappa_n(S)}$ and thus $L_r - L_{\kappa_n(r)}$ is independent of \mathcal{G} . The last inequality follows from the fact that $t-s \leq 2/n$.

Now consider the case $t \geq s+2/n$. Then we note that $r \geq s$ implies $\kappa_n(r) \geq s-(t-s)$, and in fact $\kappa_n(r) - (s-(t-s)) \geq (t-s)/2$. Recalling (3.10), we see that

$$A_{s,t} = \int_s^t (\mathcal{P}_{r-(s-(t-s))} - \mathcal{P}_{\kappa_n(r)-(s-(t-s))}) f(L_{s-(t-s)} + \mathbb{E}^{s-(t-s)} \psi_r) dr.$$

Applying (3.12) with $\rho = 0$, $\theta \wedge ((1-\varepsilon)\alpha)$ in place of β , $\mu = (\frac{\theta}{\alpha} \wedge \frac{1}{2}) + \frac{1}{2} - \varepsilon$, we get (note that all the assumptions of Proposition 3.3 are satisfied with such choice of parameters)

$$\begin{aligned}
|A_{s,t}| &\leq N \|f\|_{\mathcal{C}^\theta} \int_s^t n^{-(\frac{1}{2} + \frac{\theta}{\alpha} \wedge \frac{1}{2}) + 2\varepsilon} |t-s|^{(\frac{\theta}{\alpha} \wedge (1-\varepsilon)) - (\frac{\theta}{\alpha} \wedge \frac{1}{2}) - \frac{1}{2} + \varepsilon} dr \\
&\leq N \|f\|_{\mathcal{C}^\theta} n^{-(\frac{1}{2} + \frac{\theta}{\alpha} \wedge \frac{1}{2}) + 2\varepsilon} |t-s|^{\frac{1}{2} + \varepsilon},
\end{aligned}$$

which implies

$$\|A_{s,t}\|_{L_2(\Omega)|\mathcal{G}} \leq N \|f\|_{\mathcal{C}^\theta} n^{-(\frac{1}{2} + \frac{\theta}{\alpha} \wedge \frac{1}{2}) + 2\varepsilon} |t-s|^{\frac{1}{2} + \varepsilon}.$$

Recalling (4.27), we see that the condition (3.4) is satisfied with $\Gamma_1 = N \|f\|_{\mathcal{C}^\theta} n^{-(\frac{1}{2} + \frac{\theta}{\alpha} \wedge \frac{1}{2}) + 2\varepsilon}$ and $\varepsilon_1 = \varepsilon$.

Moving on to the condition (3.5), take $(s, t) \in \Delta_{[S, T]}$ and $u := (t+s)/2$. As before, denote $s_1 := s - (t-s)$, $s_2 := s - (u-s)$, $s_3 := s$, $s_4 := u$, $s_5 := t$. We need to bound $\mathbb{E}^{s-(t-s)} \delta A_{s,u,t} = \mathbb{E}^{s_1} \delta A_{s_3,s_4,s_5}$. By a standard computation we see that

$$\begin{aligned}
\mathbb{E}^{s_1} \delta A_{s_3,s_4,s_5} &= \mathbb{E}^{s_1} \mathbb{E}^{s_2} \int_{s_3}^{s_4} f(L_r + \mathbb{E}^{s_1} \psi_r) - f(L_{\kappa_n(r)} + \mathbb{E}^{s_1} \psi_r) \\
&\quad - f(L_r + \mathbb{E}^{s_2} \psi_r) + f(L_{\kappa_n(r)} + \mathbb{E}^{s_2} \psi_r) dr \\
&\quad + \mathbb{E}^{s_1} \mathbb{E}^{s_3} \int_{s_4}^{s_5} f(L_r + \mathbb{E}^{s_1} \psi_r) - f(L_{\kappa_n(r)} + \mathbb{E}^{s_1} \psi_r) \\
&\quad - f(L_r + \mathbb{E}^{s_3} \psi_r) + f(L_{\kappa_n(r)} + \mathbb{E}^{s_3} \psi_r) dr \\
&=: I_1 + I_2.
\end{aligned} \tag{4.28}$$

The two terms are treated in exactly the same way, so we only discuss bounding the first one. When $|t-s| \geq 4/n$, then for $r \geq s_3$ we have

$$\kappa_n(r) - s_2 \geq s_3 - 1/n - s_2 \geq (t-s)/4 > 0. \tag{4.29}$$

Therefore we first write

$$\begin{aligned}
I_1 &= \mathbb{E}^{s_1} \int_{s_3}^{s_4} (\mathcal{P}_{r-s_2} - \mathcal{P}_{\kappa_n(r)-s_2}) f(L_{s_2} + \mathbb{E}^{s_1} \psi_r) \\
&\quad - (\mathcal{P}_{r-s_2} - \mathcal{P}_{\kappa_n(r)-s_2}) f(L_{s_2} + \mathbb{E}^{s_2} \psi_r) dr.
\end{aligned}$$

Applying (3.12) with $\rho = 1$, θ in place of β , $\mu = \frac{1}{2} + \frac{\theta}{\alpha} \wedge \frac{1}{2}$, and using (4.29) yields (we see again that all the assumptions of Proposition 3.3 are satisfied with such choice of parameters)

$$\begin{aligned} |I_1| &\leq \int_{s_3}^{s_4} \|(\mathcal{P}_{r-s_2} - \mathcal{P}_{\kappa_n(r)-s_2})f\|_{\mathcal{C}^1} \mathbb{E}^{s_1} |\mathbb{E}^{s_1} \psi_r - \mathbb{E}^{s_2} \psi_r| dr \\ &\leq N \|f\|_{\mathcal{C}^\theta} \int_{s_3}^{s_4} n^{-(\frac{1}{2} + \frac{\theta}{\alpha} \wedge \frac{1}{2}) + \varepsilon} |t-s|^{-(\frac{1}{2} + \frac{\theta}{\alpha} \wedge \frac{1}{2} + \frac{1}{\alpha} - \frac{\theta}{\alpha})} |r-s_1|^\tau dr \\ &\leq N \|f\|_{\mathcal{C}^\theta} C_g n^{-(\frac{1}{2} + \frac{\theta}{\alpha} \wedge \frac{1}{2}) + \varepsilon} |t-s|^{1+\tilde{\varepsilon}}, \end{aligned} \quad (4.30)$$

where we used (4.23) in the second inequality and put $\tilde{\varepsilon} := \tau - (\frac{1}{2} + \frac{\theta}{\alpha} \wedge \frac{1}{2} + \frac{1}{\alpha} - \frac{\theta}{\alpha}) > 0$ by (4.22). By a similar argument, (4.30) holds for $|I_2|$. Therefore, taking $L_2(\Omega)|\mathcal{G}$ norm and recalling (4.28) we can conclude that in the case $|t-s| \geq 4/n$ we have

$$\|\mathbb{E}^{s-(t-s)} \delta A_{s,u,t}\|_{L_2(\Omega)|\mathcal{G}} \leq N \|f\|_{\mathcal{C}^\theta} C_g n^{-(\frac{1}{2} + \frac{\theta}{\alpha} \wedge \frac{1}{2}) + \varepsilon} |t-s|^{1+\tilde{\varepsilon}}. \quad (4.31)$$

Next, consider the case $|t-s| \leq 4/n$. By assumption, $\psi_r = g_{\kappa_n(r)}$ is $\mathcal{F}_{(\kappa_n(r)-\frac{1}{n}) \vee 0}$ measurable. Therefore, if $r \in [\kappa_n(s_1), \kappa_n(s_1) + \frac{2}{n})$, then ψ_r is $\mathcal{F}_{\kappa_n(s_1)}$ -measurable. Since $\kappa_n(s_1) \leq s_1 \leq s_2$, one has $\mathbb{E}^{s_1} \psi_r = \mathbb{E}^{s_2} \psi_r = \psi_r$ and thus the integrand in I_1 is zero. Hence, we can concentrate on the case $r \geq \kappa_n(s_1) + \frac{2}{n}$. In this case, $\kappa_n(r) - 1/n \geq s_1$. Thus we get

$$\begin{aligned} I_1 &= \mathbb{E}^{s_1} \int_{[s_3, s_4] \cap [\kappa_n(s_1) + \frac{2}{n}, 1]} \mathbb{E}^{\kappa_n(r) - \frac{1}{n}} \left(f(L_r + \mathbb{E}^{s_1} \psi_r) - f(L_{\kappa_n(r)} + \mathbb{E}^{s_1} \psi_r) \right. \\ &\quad \left. - f(L_r + \mathbb{E}^{s_2} \psi_r) + f(L_{\kappa_n(r)} + \mathbb{E}^{s_2} \psi_r) \right) dr \\ &= \mathbb{E}^{s_1} \int_{[s_3, s_4] \cap [\kappa_n(s_1) + \frac{2}{n}, 1]} \mathcal{P}_{\frac{1}{n}} f(L_{\kappa_n(r) - \frac{1}{n}} + \mathbb{E}^{s_2} \psi_r) - \mathcal{P}_{\frac{1}{n}} f(L_{\kappa_n(r) - \frac{1}{n}} + \mathbb{E}^{s_1} \psi_r) \\ &\quad + \mathcal{P}_{r - \kappa_n(r) + \frac{1}{n}} f(L_{\kappa_n(r) - \frac{1}{n}} + \mathbb{E}^{s_1} \psi_r) - \mathcal{P}_{r - \kappa_n(r) + \frac{1}{n}} f(L_{\kappa_n(r) - \frac{1}{n}} + \mathbb{E}^{s_2} \psi_r) dr, \end{aligned}$$

where we used again that ψ_r is $\mathcal{F}_{\kappa_n(r) - \frac{1}{n}}$ -measurable. Applying (3.11) with $\rho = 1$ and recalling (4.23), we immediately deduce

$$\begin{aligned} |I_1| &\leq N \|f\|_{\mathcal{C}^\theta} C_g \int_{[s_3, s_4] \cap [\kappa_n(s_1) + \frac{2}{n}, 1]} (\|\mathcal{P}_{\frac{1}{n}} f\|_{\mathcal{C}^1} + \|\mathcal{P}_{r - \kappa_n(r) + \frac{1}{n}} f\|_{\mathcal{C}^1}) \mathbb{E}^{s_1} |\mathbb{E}^{s_1} \psi_r - \mathbb{E}^{s_2} \psi_r| dr \\ &\leq N \|f\|_{\mathcal{C}^\theta} C_g |t-s|^{1+\tau} n^{-\frac{\theta-1}{\alpha}} \\ &\leq N \|f\|_{\mathcal{C}^\theta} C_g |t-s|^{1+\varepsilon} n^{-\tau - \frac{\theta-1}{\alpha} + \varepsilon} \\ &\leq N \|f\|_{\mathcal{C}^\theta} C_g |t-s|^{1+\varepsilon} n^{-(\frac{1}{2} + \frac{\theta}{\alpha} \wedge \frac{1}{2}) + \varepsilon}, \end{aligned}$$

where the penultimate inequality follows from the fact that $|t-s| \leq 4/n$, and in the last inequality we used (4.22). By the same argument, exactly the same bound holds also for $|I_2|$. Recalling now (4.28) and (4.31), we can therefore conclude that (3.5) is satisfied with $\Gamma_2 = N \|f\|_{\mathcal{C}^\theta} C_g n^{-(\frac{1}{2} + \frac{\theta}{\alpha} \wedge \frac{1}{2}) + \varepsilon}$ and $\varepsilon_2 = \varepsilon \wedge \tilde{\varepsilon}$.

It remains to verify that the process \mathcal{A}_t satisfies (3.6). Fix now $m \in \mathbb{N}$. Denote $t_i := s + i \frac{t-s}{m}$, $i = 0, \dots, m$. We get

$$\left\| \mathcal{A}_t - \mathcal{A}_s - \sum_{i=1}^{m-1} \mathcal{A}_{t_i, t_{i+1}} \right\|_{L_1(\Omega)} \leq \|\mathcal{A}_{t_1} - \mathcal{A}_{t_0}\|_{L_1(\Omega)}$$

$$\begin{aligned}
& + \left\| \sum_{i=1}^{m-1} \mathcal{A}_{t_{i+1}} - \mathcal{A}_{t_i} - \mathbb{E}^{t_i}(\mathcal{A}_{t_{i+1}} - \mathcal{A}_{t_i}) \right\|_{L_2(\Omega)} \\
& + \left\| \sum_{i=1}^{m-1} \mathbb{E}^{t_i}(\mathcal{A}_{t_{i+1}} - \mathcal{A}_{t_i}) - \mathbb{E}^{t_{i-1}}(\mathcal{A}_{t_{i+1}} - \mathcal{A}_{t_i}) \right\|_{L_2(\Omega)} \\
& + \left\| \sum_{i=1}^{m-1} A_{t_i, t_{i+1}} - \mathbb{E}^{t_{i-1}}(\mathcal{A}_{t_{i+1}} - \mathcal{A}_{t_i}) \right\|_{L_1(\Omega)} \\
& := I_1 + I_2 + I_3 + I_4.
\end{aligned} \tag{4.32}$$

Since f is bounded, we clearly have

$$I_1 \leq 2\|f\|_{C^0} m^{-1}. \tag{4.33}$$

Next, we note that the sequence $(\mathcal{A}_{t_{i+1}} - \mathcal{A}_{t_i} - \mathbb{E}^{t_i}(\mathcal{A}_{t_{i+1}} - \mathcal{A}_{t_i}))_{i=1, \dots, m-1}$ is a martingale difference sequence with respect to the filtration $(\mathcal{F}_{t_{i+1}})_{i=1, \dots, m-1}$. Therefore, the Burkholder-Davis-Gundy inequality implies

$$I_2^2 \leq \sum_{i=1}^{m-1} \left\| \mathcal{A}_{t_{i+1}} - \mathcal{A}_{t_i} - \mathbb{E}^{t_i}(\mathcal{A}_{t_{i+1}} - \mathcal{A}_{t_i}) \right\|_{L_2(\Omega)}^2 \leq 4\|f\|_{C^0}^2 m^{-1}. \tag{4.34}$$

Similarly, the sequence $(\mathbb{E}^{t_i}(\mathcal{A}_{t_{i+1}} - \mathcal{A}_{t_i}) - \mathbb{E}^{t_{i-1}}(\mathcal{A}_{t_{i+1}} - \mathcal{A}_{t_i}))_{i=1, \dots, m-1}$ is a martingale difference sequence with respect to the filtration $(\mathcal{F}_{t_i})_{i=1, \dots, m-1}$, and we get

$$I_3^2 \leq \sum_{i=1}^{m-1} \left\| \mathbb{E}^{t_i}(\mathcal{A}_{t_{i+1}} - \mathcal{A}_{t_i}) - \mathbb{E}^{t_{i-1}}(\mathcal{A}_{t_{i+1}} - \mathcal{A}_{t_i}) \right\|_{L_2(\Omega)}^2 \leq 16\|f\|_{C^0}^2 m^{-1}. \tag{4.35}$$

Finally, if $g \equiv 0$, then $I_4 = 0$. If $g \not\equiv 0$, then by condition (iii) of the theorem $\theta > 0$. Therefore, using (3.2) we derive for any $i = 1, \dots, m-1$

$$\begin{aligned}
& \left\| \mathbb{E}^{t_{i-1}}(\mathcal{A}_{t_{i+1}} - \mathcal{A}_{t_i}) - A_{t_i, t_{i+1}} \right\|_{L_1(\Omega)} \\
& \leq \left\| \int_{t_i}^{t_{i+1}} f(L_r + g_{\kappa_n(r)}) - f(L_r + \mathbb{E}^{t_{i-1}} g_{\kappa_n(r)}) \right\|_{L_1(\Omega)} \\
& \quad + \left\| \int_{t_i}^{t_{i+1}} f(L_{\kappa_n(r)} + g_{\kappa_n(r)}) + f(L_{\kappa_n(r)} + \mathbb{E}^{t_{i-1}} g_{\kappa_n(r)}) dr \right\|_{L_1(\Omega)} \\
& \leq 2\|f\|_{C^\theta} \int_{t_i}^{t_{i+1}} \|g_{\kappa_n(r)} - \mathbb{E}^{t_{i-1}} g_{\kappa_n(r)}\|_{L_1(\Omega)}^\theta dr \\
& \leq 2\|f\|_{C^\theta} (1 + C_g)(t_{i+1} - t_{i-1})^{1+\theta\tau} \leq N\|f\|_{C^\theta} (1 + C_g) m^{-1-\theta\tau},
\end{aligned}$$

where the penultimate inequality follows from the fact that if $\kappa_n(r) \leq t_{i-1}$, then $g_{\kappa_n(r)} = \mathbb{E}^{t_{i-1}} g_{\kappa_n(r)}$ and if $\kappa_n(r) \geq t_{i-1}$, then (4.23) is applicable and $\kappa_n(r) - t_{i-1} \leq t_{i+1} - t_{i-1}$. Recalling the definition of I_4 in (4.32), we get

$$I_4 \leq N\|f\|_{C^\theta} (1 + C_g) m^{-\theta\tau} \mathbb{1}_{\theta>0}. \tag{4.36}$$

Collecting together bounds (4.33)–(4.36) and substituting them into (4.32), we get

$$\left\| \mathcal{A}_t - \mathcal{A}_s - \sum_{i=1}^{m-1} A_{t_i, t_{i+1}} \right\|_{L_1(\Omega)} \leq N\|f\|_{C^\theta} (1 + C_g) m^{-\theta\tau} \mathbb{1}_{\theta>0} + N\|f\|_{C^0} m^{-\frac{1}{2}}.$$

which implies (3.6). The claimed bound (4.25) is therefore given by (3.7). \square

Remark 4.8. The reader might observe two additional constraints introduced in Lemma 4.7 compared to Lemma 4.4. Specifically, we assumed that $\mathcal{G} \subset \mathcal{F}_{\kappa_n(S)}$ and that $p = 2$. Both of these conditions are employed to derive the bound (4.27). With a generic $p > 2$, the convergence rate becomes suboptimal, given by $n^{-(\frac{1}{2} + \frac{\theta}{\alpha} \wedge \frac{1}{p}) + \varepsilon}$. Without the assumption $\mathcal{G} \subset \mathcal{F}_{\kappa_n(S)}$, there is no deterministic bound on conditional moments of $L_r - L_{\kappa_n(r)}$ given \mathcal{G} for $r \in [S, T]$.

These two restrictions substantially complicate the proof of the main results. The condition $p = 2$ is the reason why we must employ the John–Nirenberg machinery at all and bound conditional expectations, instead of merely bounding $L_p(\Omega)$ moments directly as done, for example, in [BDG21a]. The limitation $\mathcal{G} \subset \mathcal{F}_{\kappa_n(S)}$ leads to additional challenges in the buckling part of the proof of Theorem 2.2.

4.3 Bounds on $\mathcal{E}^{f,n,1}$, $\mathcal{E}^{f,n,2}$, $\mathcal{E}^{f,n,3}$

In this part of the paper we apply generic bounds from Section 4.2 to the error terms in decompositions (4.2) and (4.3).

Corollary 4.9. *Assume that all the conditions of Theorem 2.5 are satisfied. Then for any $\varepsilon > 0$ there exists a constant $N = N(\alpha, \beta, \theta, p, d, \|b\|_{\mathcal{C}^\beta}, \varepsilon, M)$ such that for any $0 \leq s \leq t \leq 1$, σ -algebra $\mathcal{G} \subset \mathcal{F}_s$ and all $n \in \mathbb{N}$ the following holds:*

$$\begin{aligned} \|\mathcal{E}_{s,t}^{f,n,1}\|_{L_p(\Omega)|\mathcal{G}} &\leq N \|f\|_{\mathcal{C}^\theta} (t-s)^{1+\frac{(\theta-1)\wedge 0}{\alpha}} \|\varphi - \varphi^n\|_{\mathcal{C}_p^0|\mathcal{G},[s,t]} \\ &\quad + N \|f\|_{\mathcal{C}^\theta} (t-s)^{\frac{3}{2}+\frac{(\theta-1)\wedge 0}{\alpha}} [\varphi - \varphi^n]_{\mathcal{C}_p^{1/2}|\mathcal{G},[s,t]}; \end{aligned} \quad (4.37)$$

$$\|\mathcal{E}_{s,t}^{f,n,2}\|_{L_p(\Omega)|\mathcal{G}} \leq N \|f\|_{\mathcal{C}^\theta} (t-s)^{1+\frac{(\theta-1)\wedge 0}{\alpha}} n^{-1} + N \|f\|_{\mathcal{C}^\theta} (t-s)^{\frac{3}{2}+\frac{(\theta-1)\wedge 0}{\alpha}} n^{-\frac{1}{2}-(\frac{\beta}{\alpha} \wedge \frac{1}{2})+\varepsilon}. \quad (4.38)$$

Proof. Recall that $\beta > 2 - 2\alpha$ and $\beta > 1 - \alpha/2$. Without loss of generality, we can assume that $\beta \leq 2$ and $\theta \leq 2$: indeed, if one of them is larger than 2 then we replace the corresponding constant by 2, this will not affect neither conditions (2.3) and (2.6), nor the bounds on the right-hand side of (4.37) and (4.38). Thus, till the end of the proof choose arbitrary $\delta > 0$ small enough so that

$$\theta \wedge \beta > \max(2 - 2\alpha, 1 - \alpha/2) + \delta\alpha. \quad (4.39)$$

To establish (4.37) and (4.38), we will apply Lemma 4.4 with $\tau = 1 + (\frac{\beta\wedge 1}{\alpha} \wedge 1) - \delta$, $\gamma = 1/2$. Thanks to (4.39), we see that condition (4.10) holds.

First, let us obtain (4.37). We apply Lemma 4.4 with $g = \varphi$, $h = \varphi^n$. It follows from (4.8) with $q = 1$ that (4.11) holds with $C_g \leq N$. Similarly, (4.9) and (3.2) imply that (4.12) holds. Thus all the assumptions of Lemma 4.4 are satisfied. Taking into account (4.7) and the fact that $1 + \frac{(\theta-1)\wedge 0}{\alpha} \leq 1 < 1 + \frac{\theta-2}{\alpha} + \tau$, we see that (4.13) yields (4.37).

Now let us prove (4.38). We will again apply Lemma 4.4 with $g_t = \varphi_t^n$, $h_t = \varphi_{\kappa_n(t)}^n$ and the same τ , γ as above. Thanks to (4.9) with $q = 1$ we see that (4.11) holds with $C_g \leq N$. Further, if $0 \leq \kappa_n(t) \leq s < t \leq 1$ then

$$\|\varphi_{\kappa_n(t)}^n - \mathbb{E}^s \varphi_{\kappa_n(t)}^n\|_{L_1(\Omega)} = 0.$$

If $0 \leq s < \kappa_n(t) \leq t \leq 1$, then applying (4.9) (in a very rough way) yields

$$\|\varphi_{\kappa_n(t)}^n - \mathbb{E}^s \varphi_{\kappa_n(t)}^n\|_{L_1(\Omega)} \leq N |\kappa_n(t) - s| \leq N(t - s).$$

Hence, (4.12) holds. Thus all the assumptions of Lemma 4.4 are satisfied. Therefore (4.13) implies

$$\begin{aligned} \|\mathcal{E}_{s,t}^{f,n,2}\|_{L_p(\Omega)|\mathcal{G}} &= \left\| \int_s^t f(L_r + \varphi_r^n) - f(L_s + \varphi_{\kappa_n(r)}^n) dr \right\|_{L_p(\Omega)|\mathcal{G}} \\ &\leq N\|f\|_{C^\theta}(t-s)^{1+\frac{(\theta-1)\wedge 0}{\alpha}} \|\varphi^n - \varphi_{\kappa_n(\cdot)}^n\|_{\mathcal{C}_p^0|\mathcal{G},[s,t]} \\ &\quad + N\|f\|_{C^\theta}(t-s)^{\frac{3}{2}+\frac{(\theta-1)\wedge 0}{\alpha}} [\varphi^n - \varphi_{\kappa_n(\cdot)}^n]_{\mathbf{C}_p^{1/2}|\mathcal{G},[s,t]}. \end{aligned} \quad (4.40)$$

Note that for any $r \in [0, 1]$, we clearly have

$$\varphi_r^n - \varphi_{\kappa_n(r)}^n = (r - \kappa_n(r))b(L_{\kappa_n(r)} + \varphi_{\kappa_n(r)}^n). \quad (4.41)$$

Since b is bounded, this implies

$$\|\varphi^n - \varphi_{\kappa_n(\cdot)}^n\|_{\mathcal{C}_p^0|\mathcal{G},[s,t]} \leq \|b\|_{C^0} \sup_{r \in [s,t]} |r - \kappa_n(r)| \leq Nn^{-1}. \quad (4.42)$$

Let now $s \leq r' \leq r \leq t$. If $\kappa_n(r) \leq r'$, then both φ_r^n and $\varphi_{\kappa_n(r)}^n$ are $\mathcal{F}_{r'}$ -measurable, so trivially

$$\varphi_r^n - \varphi_{\kappa_n(r)}^n - \mathbb{E}^{r'}[\varphi_r^n - \varphi_{\kappa_n(r)}^n] = 0 \quad (4.43)$$

Otherwise if $s \leq r' < \kappa_n(r) \leq r \leq t$, then by (4.41) and (4.6),

$$\begin{aligned} &\mathbb{E}^{r'}|\varphi_r^n - \varphi_{\kappa_n(r)}^n - \mathbb{E}^{r'}[\varphi_r^n - \varphi_{\kappa_n(r)}^n]| \\ &= (r - \kappa_n(r))\mathbb{E}^{r'}|b(L_{\kappa_n(r)} + \varphi_{\kappa_n(r)}^n) - \mathbb{E}^{r'}b(L_{\kappa_n(r)} + \varphi_{\kappa_n(r)}^n)| \\ &\leq 2(r - \kappa_n(r))\mathbb{E}^{r'}|b(L_{\kappa_n(r)} + \varphi_{\kappa_n(r)}^n) - b(L_{r'} + \varphi_{r'}^n)| \\ &\leq N(r - \kappa_n(r))\mathbb{E}^{r'}(|L_{\kappa_n(r)} - L_{r'}|^{\beta \wedge 1} \wedge 1 + |\varphi_{\kappa_n(r)}^n - \varphi_{r'}^n|^{\beta \wedge 1}) \\ &\leq N(r - \kappa_n(r))(|r - r'|^{(\frac{\beta \wedge 1}{\alpha} \wedge 1) - \varepsilon} + |r - r'|^{\beta \wedge 1}), \end{aligned}$$

where in the last inequality we used that $L_{\kappa_n(r)} - L_{r'}$ is independent of $\mathcal{F}_{r'}$, (3.13), and boundedness of b (which implies Lipschitzness of φ^n). Note that if $\alpha \geq 1$, then clearly $\beta \wedge 1 \geq \frac{\beta \wedge 1}{\alpha}$. Further, if $\alpha < 1$, then $\beta > 1/2$ thanks to (2.3). Thus in both cases $\beta \wedge 1 \geq \frac{\beta \wedge 1}{\alpha} \wedge \frac{1}{2} = \frac{\beta}{\alpha} \wedge \frac{1}{2}$, since $\alpha \leq 2$. We continue the above inequality by taking $\|\cdot\|_{L_p(\Omega)|\mathcal{G}}$ norms and writing

$$\begin{aligned} \|\mathbb{E}^{r'}|\varphi_r^n - \varphi_{\kappa_n(r)}^n - \mathbb{E}^{r'}[\varphi_r^n - \varphi_{\kappa_n(r)}^n]|\|_{L_p(\Omega)|\mathcal{G}} &\leq N(r - \kappa_n(r))|r - r'|^{(\frac{\beta}{\alpha} \wedge \frac{1}{2}) - \varepsilon} \\ &\leq Nn^{-\frac{1}{2} - (\frac{\beta}{\alpha} \wedge \frac{1}{2}) + \varepsilon} |r - r'|^{1/2}, \end{aligned}$$

where in the last inequality we used that $r - \kappa_n(r) \leq n^{-1} \wedge |r - r'|$. This together with (4.43) yields

$$[\varphi^n - \varphi_{\kappa_n(\cdot)}^n]_{\mathbf{C}_p^{1/2}|\mathcal{G},[s,t]} \leq Nn^{-\frac{1}{2} - (\frac{\beta}{\alpha} \wedge \frac{1}{2}) + \varepsilon}. \quad (4.44)$$

Substituting this and (4.42) into (4.40), we finally get

$$\|\mathcal{E}_{s,t}^{f,n,2}\|_{L_p(\Omega)} \leq N\|f\|_{C^\theta}(t-s)^{1+\frac{(\theta-1)\wedge 0}{\alpha}} n^{-1} + N\|f\|_{C^\theta}(t-s)^{\frac{3}{2}+\frac{(\theta-1)\wedge 0}{\alpha}} n^{-\frac{1}{2} - (\frac{\beta}{\alpha} \wedge \frac{1}{2}) + \varepsilon}. \quad \square$$

Remark 4.10. We see now why it was important in the derivation of (4.38) that the seminorm $[\cdot]_{\mathbf{C}_p^\gamma|\mathcal{G},[s,t]}$ rather than $[\cdot]_{\mathcal{C}_p^\gamma|\mathcal{G},[s,t]}$ appeared in (4.13), recall Remark 4.5. Indeed, by taking $r \in [s, t]$ such that $r = \kappa_n(r)$ and $s \leq r' \leq r \leq t$ it is easy to see that one has

$$\|\varphi_r^n - \varphi_{\kappa_n(r)}^n - \varphi_{r'}^n - \varphi_{\kappa_n(r')}^n\|_{L_p(\Omega)|\mathcal{G}} = (r' - \kappa_n(r'))\|b(L_{\kappa_n(r')} + \varphi_{\kappa_n(r')}^n)\|_{L_p(\Omega)|\mathcal{G}}.$$

This implies that $[\varphi^n - \varphi_{\kappa_n(\cdot)}^n]_{\mathcal{C}_p^\gamma|\mathcal{G},[s,t]} = \infty$ for any $\gamma > 0$, and is obviously much worse than (4.44). Note however that for (4.37) we simply bounded $[\cdot]_{\mathbf{C}_p^\gamma|\mathcal{G},[s,t]}$ by $[\cdot]_{\mathcal{C}_p^\gamma|\mathcal{G},[s,t]}$.

Corollary 4.11. *Assume that all the conditions of Theorem 2.5 are satisfied. Then for any $\varepsilon \in (0, 1/2)$ there exists a constant $N = N(\alpha, \beta, \theta, d, \|b\|_{\mathcal{C}^\beta}, \varepsilon, M)$ such that for any $0 \leq s \leq t \leq 1$, $n \in \mathbb{N}$, and any σ -algebra $\mathcal{G} \subset \mathcal{F}_{\kappa_n(s)}$ the following holds:*

$$\|\mathcal{E}_{s,t}^{f,n,3}\|_{L_2(\Omega)|\mathcal{G}} \leq N \|f\|_{\mathcal{C}^\theta} n^{-(\frac{1}{2} + \frac{\theta}{\alpha} \wedge \frac{1}{2}) + 2\varepsilon} |t - s|^{\frac{1}{2} + \varepsilon}. \quad (4.45)$$

Proof. Choose $\delta > 0$ small enough so that

$$\theta \wedge \beta \wedge 1 > 1 - \frac{\alpha}{2} + \delta\alpha.$$

We apply Lemma 4.7 with $g = \varphi^n$ and $\theta \wedge 1$ in place of θ . Note that the rate provided in (4.25) is consistent with (4.45), since $\frac{\theta}{\alpha} \wedge \frac{1}{2} = \frac{\theta \wedge 1}{\alpha} \wedge \frac{1}{2}$. Therefore it remains to verify the conditions of Lemma 4.7.

For any $0 \leq s \leq t \leq 1$, $r \in [s, 1]$ we have by (4.9) with $q = 1$

$$\mathbb{E}^s |\mathbb{E}^s \varphi_r^n - \mathbb{E}^t \varphi_r^n| = \mathbb{E}^s |\mathbb{E}^t [\mathbb{E}^s \varphi_r^n - \varphi_r^n]| \leq \mathbb{E}^s |\mathbb{E}^s \varphi_r^n - \varphi_r^n| \leq N |r - s|^{1 - \delta + \frac{\beta \wedge \alpha \wedge 1}{\alpha}}.$$

Therefore, condition (4.23) is satisfied with $\tau = 1 - \delta + \frac{\beta \wedge \alpha \wedge 1}{\alpha}$ and $C_g = N$. We note that if $\beta \wedge \alpha \wedge 1 = \beta \wedge 1$, then using that $\frac{\theta}{\alpha} \wedge \frac{1}{2} \leq \frac{\theta \wedge 1}{\alpha}$ and $\frac{1}{2} - \frac{1}{\alpha} + \frac{\beta \wedge 1}{\alpha} > \delta$, we get

$$\tau = 1 - \delta + \frac{\beta \wedge 1}{\alpha} > \frac{1}{2} + \frac{1}{\alpha} \geq \frac{1}{2} + \frac{1}{\alpha} + \frac{\theta}{\alpha} \wedge \frac{1}{2} - \frac{\theta \wedge 1}{\alpha}$$

and condition (4.22) holds. Alternatively, if $\beta \wedge \alpha \wedge 1 = \alpha$, then we use that $\frac{1}{2} \geq \frac{\theta}{\alpha} \wedge \frac{1}{2}$ and $1 - \frac{1}{\alpha} + \frac{\theta \wedge 1}{\alpha} > \delta$ to get

$$\tau = 2 - \delta > \frac{1}{2} + \frac{1}{\alpha} + \frac{\theta}{\alpha} \wedge \frac{1}{2} - \frac{\theta \wedge 1}{\alpha}.$$

Hence also in this case condition (4.22). We also see that $\varphi_{\kappa_n(t)}^n$ is $\mathcal{F}_{(\kappa_n(t) - \frac{1}{n}) \vee 0}$ -measurable for $t \in [0, 1]$. Thus all the conditions of Lemma 4.7 are satisfied and we get (4.45). \square

Gathering the error bounds of Corollary 4.9 and Corollary 4.11, we finally derive the following crucial conditional quadrature estimate.

Corollary 4.12. *Assume that all the conditions of Theorem 2.5 are satisfied. Then for any $\varepsilon \in (0, \frac{1}{2} + \frac{(\theta-1) \wedge 0}{\alpha})$ there exists a constant $N_0 = N_0(\alpha, \beta, \theta, d, \|b\|_{\mathcal{C}^\beta}, \varepsilon, M)$ such that for any $n \in \mathbb{N}$, $k \in \{0, 1, \dots, n\}$, $0 \leq k/n \leq s \leq t \leq 1$, the following holds:*

$$\begin{aligned}
& \left\| \int_s^t f(X_r) dr - \int_s^t f(X_{\kappa_n(r)}^n) dr \right\|_{L_2(\Omega) | \mathcal{F}_{\frac{k}{n}}} \\
& \leq N_0 \|f\|_{C^\theta} |t-s|^{\frac{1}{2}+\varepsilon} (n^{-(\frac{1}{2}+\frac{\beta \wedge \theta}{\alpha} \wedge \frac{1}{2})+2\varepsilon} + \|\varphi - \varphi^n\|_{\mathcal{C}_2^0 | \mathcal{F}_{\frac{k}{n}}[s,t]} + [\varphi - \varphi^n]_{\mathcal{C}_2^{1/2} | \mathcal{F}_{\frac{k}{n}}[s,t]}).
\end{aligned} \tag{4.46}$$

Proof. Fix $k/n \leq s \leq t \leq 1$. We use (4.37), (4.38), and (4.45) with $\mathcal{G} = \mathcal{F}_{\frac{k}{n}}$ to bound each term in (4.2). Since $k/n \leq s$, we see that $k/n \leq \kappa_n(s)$. Therefore, $\mathcal{F}_{\frac{k}{n}} \subset \mathcal{F}_{\kappa_n(s)}$ and the conditions of Corollary 4.11 hold. Using that $1 + \frac{(\theta-1) \wedge 0}{\alpha} > \frac{1}{2} + \varepsilon$ and $(\frac{1}{2} + \frac{\beta \wedge \theta}{\alpha} \wedge \frac{1}{2}) - \varepsilon < 1$, we get the desired bound. \square

4.4 Main proofs

Proof of Theorem 2.2. Fix $\varepsilon > 0$ small enough so that

$$\gamma_0 := \frac{1}{2} + \frac{\beta}{\alpha} \wedge \frac{1}{2} - 2\varepsilon > 0.$$

In the proof we will apply (4.46) with $f = b$, $\theta = \beta$. We take now $\Delta > 0$ small enough so that

$$\Delta^\varepsilon N_0 \|b\|_{C^\beta} \leq 1/4, \tag{4.47}$$

where N_0 is the constant from (4.46). Recall the decomposition (4.1). Fix $n \in \mathbb{N}$ and put

$$\mathcal{A}_t^n := X_t - X_t^n = \varphi_t - \varphi_t^n = \int_0^t (b(X_r) - b(X_{\kappa_n(r)}^n)) dr + x_0 - x_0^n, \quad t \in [0, 1]. \tag{4.48}$$

Step 1. We claim that for any $S \in \{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\}$, $T \in [S, 1 \wedge (S + \Delta)]$ we have

$$[\mathcal{A}^n]_{\mathcal{C}_2^{1/2} | \mathcal{F}_{S,[S,T]}} \leq n^{-\gamma_0} + |\mathcal{A}_S^n|. \tag{4.49}$$

Indeed, fix $S \in \{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\}$, $T \in [S, 1 \wedge (S + \Delta)]$. Note that for any $S \leq s \leq t \leq T$

$$\mathcal{A}_t^n - \mathcal{A}_s^n = (\varphi_t - \varphi_s) - (\varphi_t^n - \varphi_s^n) = \int_s^t b(X_r) dr - \int_s^t b(X_{\kappa_n(r)}^n) dr.$$

By taking in (4.46) $k/n = S$, we get from the above identity

$$\begin{aligned}
\|\mathcal{A}_t^n - \mathcal{A}_s^n\|_{L_2(\Omega) | \mathcal{F}_S} & \leq N_0 \|b\|_{C^\beta} |t-s|^{\frac{1}{2}+\varepsilon} (n^{-\gamma_0} + \|\varphi - \varphi^n\|_{\mathcal{C}_2^0 | \mathcal{F}_{S,[S,T]}} + [\varphi - \varphi^n]_{\mathcal{C}_2^{1/2} | \mathcal{F}_{S,[S,T]}}) \\
& \leq N_0 \|b\|_{C^\beta} |t-s|^{\frac{1}{2}+\varepsilon} (n^{-\gamma_0} + |\varphi_S - \varphi_S^n| + 2[\varphi - \varphi^n]_{\mathcal{C}_2^{1/2} | \mathcal{F}_{S,[S,T]}}),
\end{aligned} \tag{4.50}$$

where we used that if $s, t \in [S, T]$, then trivially both (semi)norms on the right-hand side of equation (4.46) can be replaced by the ones on $[S, T]$. Dividing (4.50) by $|t-s|^{1/2}$ and taking supremum over $s, t \in [S, T]$, $s \leq t$, we get

$$[\mathcal{A}^n]_{\mathcal{C}_2^{1/2} | \mathcal{F}_{S,[S,T]}} \leq N_0 \|b\|_{C^\beta} \Delta^\varepsilon (n^{-\gamma_0} + |\mathcal{A}_S^n| + 2[\mathcal{A}^n]_{\mathcal{C}_2^{1/2} | \mathcal{F}_{S,[S,T]}}),$$

where we also used the definition of \mathcal{A}^n in (4.48) and the inequality $t - s \leq T - S \leq \Delta$. By our choice of Δ (4.47), the estimate buckles and by putting $2N_0\|b\|_{C^\beta}\Delta^\varepsilon[\mathcal{A}^n]_{\mathcal{C}_2^{1/2}|\mathcal{F}_S, [S, T]}$ to the left-hand side, we get (4.49).

Step 2. We get rid of the assumption that S is a gridpoint and claim that for any $0 \leq S < T \leq 1$, $T \in [S, 1 \wedge (S + \Delta)]$ we have

$$\|\mathcal{A}_T^n - \mathcal{A}_S^n\|_{L_2(\Omega)|\mathcal{F}_S} \leq \Delta^{\frac{1}{2}}(n^{-\gamma_0} + |\mathcal{A}_S^n|) + 4\|b\|_{C^0}n^{-1}. \quad (4.51)$$

Let $0 \leq S < T \leq 1$, $T \leq S + \Delta$. If $T - S \leq \frac{1}{n}$, there is nothing to prove: (4.51) follows from the fact that $|\mathcal{A}_T^n - \mathcal{A}_S^n| \leq 2\|b\|_{C^0}|T - S| \leq 2\|b\|_{C^0}n^{-1}$. Therefore, we assume from now on that $T > S + \frac{1}{n}$. Denote $S' := \kappa_n(S) + \frac{1}{n}$; that is, S' is the smallest gridpoint strictly bigger than S . We see that $S' > S$ and therefore we have

$$\begin{aligned} \|\mathcal{A}_T^n - \mathcal{A}_S^n\|_{L_2(\Omega)|\mathcal{F}_S} &\leq \|\mathcal{A}_T^n - \mathcal{A}_{S'}^n\|_{L_2(\Omega)|\mathcal{F}_S} + \|\mathcal{A}_{S'}^n - \mathcal{A}_S^n\|_{L_2(\Omega)|\mathcal{F}_S} \\ &= (\mathbb{E}[\mathbb{E}(|\mathcal{A}_T^n - \mathcal{A}_{S'}^n|^2|\mathcal{F}_{S'})|\mathcal{F}_S])^{\frac{1}{2}} + \|\mathcal{A}_{S'}^n - \mathcal{A}_S^n\|_{L_2(\Omega)|\mathcal{F}_S}. \end{aligned} \quad (4.52)$$

Next, since S' is a gridpoint we have $S' \in \{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\}$ and $T \leq S' + \Delta$. Therefore, the results of Step 1 are applicable and we get from (4.49)

$$\|\mathcal{A}_T^n - \mathcal{A}_{S'}^n\|_{L_2(\Omega)|\mathcal{F}_{S'}} \leq \Delta^{\frac{1}{2}}(n^{-\gamma_0} + |\mathcal{A}_{S'}^n|) \leq \Delta^{\frac{1}{2}}(n^{-\gamma_0} + |\mathcal{A}_S^n|) + |\mathcal{A}_S^n - \mathcal{A}_{S'}^n|.$$

Substituting this back into (4.52) and using again that $|\mathcal{A}_S^n - \mathcal{A}_{S'}^n| \leq 2\|b\|_{C^0}|S - S'| \leq 2\|b\|_{C^0}n^{-1}$, we get (4.51).

Step 3. Now with (4.51) in hand we apply the weighted John-Nirenberg inequality Proposition 3.2 to the process \mathcal{A}^n introduced in (4.48). Setting $\xi_t^n := \Delta^{\frac{1}{2}}(n^{-\gamma_0} + |\mathcal{A}_t^n|) + 4\|b\|_{C^0}n^{-1}$, $t \in [0, 1]$, we see that both processes \mathcal{A}^n and ξ^n are continuous and (3.8) holds thanks to (4.51). Therefore (3.9) implies that for any $p \geq 1$ there exists $N_1 = N_1(p)$ independent of n such that for any $0 \leq S < T \leq 1$, $T \leq S + \Delta$

$$\begin{aligned} \left\| \sup_{r \in [S, T]} |\mathcal{A}_r^n - \mathcal{A}_S^n| \right\|_{L_p(\Omega)|\mathcal{F}_S} &\leq N_1\Delta^{\frac{1}{2}}n^{-\gamma_0} + N_1\Delta^{\frac{1}{2}} \left\| \sup_{r \in [S, T]} |\mathcal{A}_r^n| \right\|_{L_p(\Omega)|\mathcal{F}_S} + 4N_1\|b\|_{C^0}n^{-1} \\ &\leq N_1\Delta^{\frac{1}{2}}n^{-\gamma_0} + N_1\Delta^{\frac{1}{2}}|A_S| \\ &\quad + N_1\Delta^{\frac{1}{2}} \left\| \sup_{r \in [S, T]} |\mathcal{A}_r^n - A_S| \right\|_{L_p(\Omega)|\mathcal{F}_S} + 4N_1\|b\|_{C^0}n^{-1}. \end{aligned}$$

Take now Δ small enough so that in addition to our standing assumption (4.47) we have also

$$N_1\Delta^{\frac{1}{2}} \leq 1/2.$$

Then the estimate buckles and we get

$$\left\| \sup_{r \in [S, T]} |\mathcal{A}_r^n - \mathcal{A}_S^n| \right\|_{L_p(\Omega)} \leq n^{-\gamma_0} + \|A_S\|_{L_p(\Omega)} + 8N_1\|b\|_{C^0}n^{-1},$$

whenever $0 \leq S < T \leq 1$, $T - S < \Delta$. Iterating this bound $\lceil \Delta^{-1} \rceil$ times yields

$$\left\| \sup_{r \in [0, 1]} |\mathcal{A}_r^n| \right\|_{L_p(\Omega)} \leq Nn^{-\gamma_0} + N|x_0 - x_0^n|,$$

for some $N = N(d, \alpha, \beta, p, \varepsilon, M, \|b\|_{C^\beta})$, which is the claimed error estimate (2.4). \square

Proof of Corollary 2.3. The statement follows from Theorem 2.2 by the standard argument. Namely, we fix $\varepsilon > 0$ and choose $p > 1/\varepsilon$. Set

$$\eta(\omega) := \sup_{n \in \mathbb{N}} (n^{\frac{1}{2} + \frac{\beta}{\alpha} \wedge \frac{1}{2} - 2\varepsilon} \|X(\omega) - X^n(\omega)\|_{C^0([0,1])}).$$

Then by Theorem 2.2,

$$\begin{aligned} \mathbb{E}\eta^p &= \mathbb{E}[\sup_{n \in \mathbb{N}} (n^{\frac{1}{2} + \frac{\beta}{\alpha} \wedge \frac{1}{2} - 2\varepsilon} \|X(\omega) - X^n(\omega)\|_{C^0([0,1])})^p] \\ &\leq \mathbb{E}[\sum_{n \in \mathbb{N}} n^{p(\frac{1}{2} + \frac{\beta}{\alpha} \wedge \frac{1}{2} - 2\varepsilon)} \|X(\omega) - X^n(\omega)\|_{C^0([0,1])}^p] \\ &\leq N \sum_{n \in \mathbb{N}} n^{-p\varepsilon} < \infty. \end{aligned}$$

Thus $\eta < \infty$ a.s. which completes the proof. \square

Proof of Theorem 2.5. We argue as in the proof of Corollary 4.12, but we skip the buckling step. Take $\varepsilon > 0$ small enough so that

$$\gamma_f := \frac{1}{2} + \frac{\beta \wedge \theta}{\alpha} \wedge \frac{1}{2} - 2\varepsilon > 0.$$

Fix $n \in \mathbb{N}$ and put

$$\mathcal{A}_t^{n,f} := \int_0^t (f(X_r) - f(X_{\kappa_n(r)}^n)) dr, \quad t \in [0, 1]. \quad (4.53)$$

Step 1. Take $\Delta = \Delta(\alpha, \beta, d, \|b\|_{C^\beta}, \varepsilon, M)$ as in (4.47). Then Corollary 4.12 and (4.49) imply for any $S \in \{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\}$, $T \in [S, 1 \wedge (S + \Delta)]$

$$\begin{aligned} \|\mathcal{A}_T^{n,f} - \mathcal{A}_S^{n,f}\|_{L_2(\Omega)|\mathcal{F}_S} &\leq N\|f\|_{C^\theta}(n^{-\gamma_f} + \|\varphi - \varphi^n\|_{\mathcal{C}_2^0|\mathcal{F}_S, [S, T]} + [\varphi - \varphi^n]_{\mathcal{C}_2^{1/2}|\mathcal{F}_S, [S, T]}) \\ &\leq N\|f\|_{C^\theta}(n^{-\gamma_f} + |\varphi_S - \varphi_S^n| + 2[\varphi - \varphi^n]_{\mathcal{C}_2^{1/2}|\mathcal{F}_S, [S, T]}) \\ &\leq N\|f\|_{C^\theta}(n^{-\gamma_f} + |\varphi_S - \varphi_S^n|) \end{aligned} \quad (4.54)$$

for $N = N(\alpha, \beta, \theta, d, \|b\|_{C^\beta}, \varepsilon, M)$.

Step 2. We remove the restriction in (4.54) that S is a grid point. Let $0 \leq S < T \leq 1$, $T \leq S + \Delta$. We note that again that if $T - S \leq \frac{1}{n}$, then $|\mathcal{A}_T^{n,f} - \mathcal{A}_S^{n,f}| \leq 2\|f\|_{C^0} n^{-1}$ and (4.54) holds. Otherwise, if $T > S + \frac{1}{n}$, we put $S' := \kappa_n(S) + \frac{1}{n}$. Since S' is a gridpoint, we have from Step 1

$$\begin{aligned} \|\mathcal{A}_T^{n,f} - \mathcal{A}_{S'}^{n,f}\|_{L_2(\Omega)|\mathcal{F}_{S'}} &\leq N\|f\|_{C^\theta}(n^{-\gamma_f} + |\varphi_{S'} - \varphi_{S'}^n|) \\ &\leq N\|f\|_{C^\theta}(n^{-\gamma_f} + |\varphi_S - \varphi_S^n|) + N\|f\|_{C^0} n^{-1}, \end{aligned} \quad (4.55)$$

where we used that $|\varphi_S - \varphi_{S'}| \leq \|b\|_{C^0} n^{-1}$ and $|\varphi_S^n - \varphi_{S'}^n| \leq \|b\|_{C^0} n^{-1}$. Therefore, using that $\mathcal{F}_S \subset \mathcal{F}_{S'}$ and the boundedness of f , we get from (4.55) for any $0 \leq S < T \leq 1$, $T \leq S + \Delta$

$$\|\mathcal{A}_T^{n,f} - \mathcal{A}_S^{n,f}\|_{L_2(\Omega)|\mathcal{F}_S} \leq \|\mathcal{A}_T^{n,f} - \mathcal{A}_{S'}^{n,f}\|_{L_2(\Omega)|\mathcal{F}_S} + \|\mathcal{A}_{S'}^{n,f} - \mathcal{A}_S^{n,f}\|_{L_2(\Omega)|\mathcal{F}_S}$$

$$\leq N\|f\|_{C^\theta}(n^{-\gamma_f} + |\varphi_S - \varphi_S^n|), \quad (4.56)$$

where $N = N(\alpha, \beta, \theta, d, \|b\|_{C^\beta}, \varepsilon, M)$.

Step 3. Now we apply the weighted John-Nirenberg inequality Proposition 3.2 to the process $\mathcal{A}^{n,f}$. Thanks to the boundedness of f and b the processes $\mathcal{A}^{n,f}$ and $\varphi - \varphi^n$ are continuous. Therefore, by Proposition 3.2, inequality (4.56) yields for any $0 \leq S < T \leq 1$, $T \leq S + \Delta$, $p \geq 1$

$$\begin{aligned} \left\| \sup_{r \in [S, T]} |\mathcal{A}_r^{n,f} - \mathcal{A}_S^{n,f}| \right\|_{L_p(\Omega)} &\leq N\|f\|_{C^\theta} n^{-\gamma_f} + N\|f\|_{C^\theta} \left\| \sup_{r \in [S, T]} |\varphi_r - \varphi_r^n| \right\|_{L_p(\Omega)} \\ &\leq N\|f\|_{C^\theta}(n^{-\gamma_f} + |x_0 - x_0^n|), \end{aligned}$$

where the last inequality follows from Theorem 2.2 and $N = N(\alpha, \beta, \theta, d, \|b\|_{C^\beta}, \varepsilon, M, p)$ is independent of S, T . Iterating the above bound $\lceil \Delta^{-1} \rceil$ times, we get (2.5). \square

Proof of Theorem 2.4. We will use the same proof strategy as in the proofs of Theorems 2.2 and 2.5: namely, we first obtain the conditional bound at grid points, then remove the grid points restriction, and finally, apply the John-Nirenberg inequality.

We note that the theorem does not follow directly from Theorem 2.5 because the function f is allowed to be of any non-negative regularity (or just bounded if $\theta = 0$). We stress that we do not impose the restriction $\theta > 1 - \frac{\alpha}{2}$ as before.

Take $\varepsilon > 0$ small enough so that

$$\gamma_{L,f} := \frac{1}{2} + \frac{\theta}{\alpha} \wedge \frac{1}{2} - 2\varepsilon > 0.$$

Fix $n \in \mathbb{N}$ and put

$$\mathcal{A}_t^{n,L,f} := \int_0^t (f(L_r) - f(L_{\kappa_n(r)}^n)) dr, \quad t \in [0, 1].$$

Step 1. Let $S \in \{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\}$, $T \in [S, 1]$ and apply Lemma 4.7 with $g \equiv 0$. We get

$$\|\mathcal{A}_T^{n,L,f} - \mathcal{A}_S^{n,L,f}\|_{L_2(\Omega)|\mathcal{F}_S} \leq N\|f\|_{C^\theta} n^{-\gamma_{L,f}} \quad (4.57)$$

for $N = N(\alpha, \theta, d, \varepsilon, M)$.

Step 2. Now we take arbitrary $S \in [0, 1]$. Let $T \in [S, 1]$. Let $S' := \kappa_n(S) + \frac{1}{n}$. As before, if $T - S \leq 1/n$, then there is nothing to prove: one have $|\mathcal{A}_T^{n,L,f} - \mathcal{A}_S^{n,L,f}| \leq 2\|f\|_{C^0} n^{-1}$. Alternatively, we have $T \geq S'$ and therefore by (4.57) and boundedness of f

$$\begin{aligned} \|\mathcal{A}_T^{n,L,f} - \mathcal{A}_S^{n,L,f}\|_{L_2(\Omega)|\mathcal{F}_{S'}} &\leq \|\mathcal{A}_T^{n,L,f} - \mathcal{A}_{S'}^{n,L,f}\|_{L_2(\Omega)|\mathcal{F}_{S'}} + \|\mathcal{A}_{S'}^{n,L,f} - \mathcal{A}_S^{n,L,f}\|_{L_2(\Omega)|\mathcal{F}_{S'}} \\ &\leq N\|f\|_{C^\theta} n^{-\gamma_{L,f}}. \end{aligned}$$

This implies that (4.57) holds for any $0 \leq S \leq T \leq 1$.

Step 3. Since the process $\mathcal{A}^{n,L,f}$ is continuous and (4.57) holds for any $0 \leq S \leq T \leq 1$, we see that all the conditions of the John-Nirenberg inequality Proposition 3.2 are satisfied. Therefore for any $p \geq 1$ there exists a constant $N = N(\alpha, \theta, d, \varepsilon, p, M) > 0$ such that

$$\left\| \sup_{r \in [0, 1]} |\mathcal{A}_r^{n,L,f}| \right\|_{L_p(\Omega)} \leq N\|f\|_{C^\theta} n^{-\gamma_{L,f}},$$

which is the desired bound (2.5). \square

Appendix A. Proofs of the well-posedness of SDE (1.1)

To show that SDE (1.1) is strongly well-posed we will run a fixed point argument. Throughout the Appendix we will assume that assumptions H1–H3 holds. We will assume without loss of generality that $\beta \leq 2$. For a measurable bounded function $f: [0, 1] \times \Omega \rightarrow \mathbb{R}^d$, $\gamma \in (0, 1]$, $p \geq 1$, $T \in (0, 1]$ put

$$\|f\|_{\mathcal{C}_p^\gamma, [0, T]} := \sup_{t \in [0, T]} \|f(t)\|_{L_p(\Omega)} + \sup_{s, t \in [0, T]} \frac{\|f(t) - f(s)\|_{L_p(\Omega)}}{(t - s)^\gamma}.$$

and consider a mapping

$$\mathcal{S}f(t) := \eta + \int_0^t b(L_s + f_s) ds, \quad 0 \leq t \leq 1,$$

where $\eta \in \mathbb{R}^d$ is a \mathcal{F}_0 -measurable vector. We claim the following contraction bound.

Lemma A.1. *Let $\varphi, \psi: [0, 1] \times \Omega \rightarrow \mathbb{R}^d$ be bounded, adapted, measurable functions. Suppose that $1 - \alpha/2 < \beta < 2$. Assume that there exist constants $\tau \in (0, 1]$, $\tau > (2 - \beta)/\alpha$, $C_\varphi, C_\psi > 0$, $T_0 \in (0, 1]$, $S \in [0, 1 - T_0]$ such that for any $S \leq s \leq t \leq S + T_0$ one has a.s.*

$$\mathbb{E}^s |\varphi_t - \mathbb{E}^s \varphi_t| \leq C_\varphi |t - s|^\tau; \quad \mathbb{E}^s |\psi_t - \mathbb{E}^s \psi_t| \leq C_\psi |t - s|^\tau. \quad (\text{A.1})$$

Then for any $p \geq 2$ there exist constants $N = N(\alpha, \beta, \tau, p, d, \|b\|_{\mathcal{C}^\beta})$, $\varepsilon = \varepsilon(\alpha, \beta, \tau) > 0$ such that for any $T \in (0, T_0]$ and $S \in [0, 1 - T]$ one has

$$\|\mathcal{S}\varphi - \mathcal{S}\psi\|_{\mathcal{C}_p^{\frac{1}{2}}, [S, S+T]} \leq N(1 + C_\varphi)T^\varepsilon \|\varphi - \psi\|_{\mathcal{C}_p^{\frac{1}{2}}, [S, S+T]} + \|\mathcal{S}\varphi(S) - \mathcal{S}\psi(S)\|_{L_p(\Omega)}. \quad (\text{A.2})$$

Proof. Fix $T \in (0, T_0]$. We will apply Lemma 4.4 with $g = \varphi$, $h = \psi$, $\gamma = 1/2$, $\theta = \beta$, $f = b$, and $\mathcal{G} = \{\emptyset, \Omega\}$. We see that all the conditions of the Lemma are satisfied and it follows from (4.13) that for any $S \leq s \leq t \leq S + T$

$$\begin{aligned} \|(\mathcal{S}\varphi(t) - \mathcal{S}\psi(t)) - (\mathcal{S}\varphi(s) - \mathcal{S}\psi(s))\|_{L_p(\Omega)} &= \left\| \int_s^t (b(L_s + \varphi_s) - b(L_s + \psi_s)) ds \right\|_{L_p(\Omega)} \\ &\leq N(1 + C_\varphi) |t - s|^{\frac{1}{2} + \varepsilon} \|\varphi - \psi\|_{\mathcal{C}_p^{\frac{1}{2}}, [0, T]}, \end{aligned} \quad (\text{A.3})$$

where $\varepsilon := \frac{1}{2} + \frac{(\beta-1) \wedge 0}{\alpha} > 0$ thanks to (2.3). Therefore, by taking in the above inequality $s = S$, we get

$$\|\mathcal{S}\varphi(t) - \mathcal{S}\psi(t)\|_{L_p(\Omega)} \leq NT^\varepsilon (1 + C_\varphi) \|\varphi - \psi\|_{\mathcal{C}_p^{\frac{1}{2}}, [0, T]} + \|\mathcal{S}\varphi(S) - \mathcal{S}\psi(S)\|_{L_p(\Omega)}.$$

Combining this with (A.3), we get the desired bound (A.2). \square

Proof of Theorem 2.1: strong uniqueness. Let X, \tilde{X} be two strong solutions to (1.1) with the initial condition η adapted to the same filtration \mathcal{F} ; here η is \mathcal{F}_0 -measurable vector in \mathbb{R}^d . Define $\varphi := X - L$, $\tilde{\varphi} := \tilde{X} - L$. Since X, \tilde{X} solve (1.1), we obviously have $\mathcal{S}\varphi = \varphi$, $\mathcal{S}\tilde{\varphi} = \tilde{\varphi}$.

It follows from (2.3) that one can choose $\delta > 0$ small enough so that

$$\beta > \max(2 - 2\alpha + \delta\alpha, 1 - \alpha/2 + \delta\alpha). \quad (\text{A.4})$$

We apply Lemma A.1 with $S = 0$ to the functions $\varphi, \tilde{\varphi}$. By Lemma 4.2(iii), condition (A.1) holds with $T_0 = 1, \tau = 1 + (\frac{\beta \wedge \alpha \wedge 1}{\alpha}) - \delta$ and $C_\varphi = C_\psi = N$. It is easy to see that with such choice of τ , one has $\tau > (2 - \beta)/\alpha$, and thus all the conditions of Lemma A.1 hold. Therefore, (A.2) implies for any $T \in (0, 1]$

$$\|\varphi - \tilde{\varphi}\|_{\mathcal{C}_2^{\frac{1}{2}}, [0, T]} = \|\mathcal{S}\varphi - \mathcal{S}\tilde{\varphi}\|_{\mathcal{C}_2^{\frac{1}{2}}, [0, T]} \leq NT^\varepsilon \|\varphi - \tilde{\varphi}\|_{\mathcal{C}_2^{\frac{1}{2}}, [0, T]}. \quad (\text{A.5})$$

Note that $\|\varphi - \tilde{\varphi}\|_{\mathcal{C}_2^{\frac{1}{2}}, [0, T]} < \infty$ regardless of the moment conditions imposed on η (we did not impose any). Indeed,

$$\|\varphi - \tilde{\varphi}\|_{\mathcal{C}_2^{\frac{1}{2}}, [0, T]} = \left\| \int_0^\cdot (b(L_s + \varphi_s) - b(L_s + \tilde{\varphi}_s)) ds \right\|_{\mathcal{C}_2^{\frac{1}{2}}, [0, T]} \leq 2\|b\|_{\mathcal{C}^0}.$$

Since N in (A.5) does not depend on T , by taking in (A.5) T small enough such that $NT^\varepsilon \leq 1/2$, we have $\varphi = \tilde{\varphi}$ on $[0, T]$, which implies $X = \tilde{X}$ on $[0, T]$. Repeating this procedure $\lceil 1/T \rceil$ times by starting at time iT instead of 0, $i = 1, \dots, \lceil 1/T \rceil - 1$, (note that we have not imposed any moment conditions on η , so there is no problem that, e.g., X_T does not have a second moment) we get strong uniqueness on the interval $[0, 1]$. \square

To establish strong existence, we consider a sequence $\varphi^{(0)}(t) := \eta, 0 \leq t \leq 1, \varphi^{(n)} := \mathcal{S}^n \varphi^{(0)}, n \in \mathbb{N}$. Clearly, for any $n \in \mathbb{Z}_+, t \in [0, 1]$

$$\varphi^{(n+1)}(t) = \eta + \int_0^t b(L_s + \varphi^{(n)}(s)) ds. \quad (\text{A.6})$$

Lemma A.2. *For any $\varepsilon > 0$, there exist constants $N = N(\alpha, \beta, \varepsilon, \|b\|_{\mathcal{C}^\beta}) > 0, T_0 = T_0(\alpha, \beta, \varepsilon, \|b\|_{\mathcal{C}^\beta}) \in (0, 1]$ such that for any $n \in \mathbb{Z}_+, S \leq s \leq t \leq S + T_0$*

$$\mathbb{E}^s |\varphi_t^{(n)} - \mathbb{E}^s \varphi_t^{(n)}| \leq N|t - s|^{1 + \frac{\beta \wedge 1 \wedge \alpha}{\alpha} - \varepsilon}. \quad (\text{A.7})$$

Proof. Fix $\varepsilon > 0$. Denote $m := 1 + \frac{\beta \wedge 1 \wedge \alpha}{\alpha} - \varepsilon$. First we note that

$$1 + (\beta \wedge 1)m - m > 0. \quad (\text{A.8})$$

Indeed, if $\beta \geq 1$ this is obvious; otherwise since $\beta > 1 - \alpha$ we have

$$m < 1 + \frac{\beta}{\alpha} < 1 + \frac{\beta}{1 - \beta} = \frac{1}{1 - \beta},$$

which implies (A.8). Fix now $T_0 \in (0, 1]$ small enough such that

$$(8\|b\|_{\mathcal{C}^\beta})^{\beta \wedge 1} T_0^{1 + (\beta \wedge 1)m - m} \leq 1. \quad (\text{A.9})$$

We will prove by induction over n that (A.7) holds with $8M^*\|b\|_{\mathcal{C}^\beta}$ in place of N , where

$$M^* := M(\beta \wedge (\alpha - \varepsilon\alpha/2), \varepsilon/2) + 1;$$

recall the definition of the function $M = M(p, \varepsilon)$ in H3.

The case $n = 0$ is obvious. Assume that the lemma holds for some $n \in \mathbb{Z}_+$, and let us prove it for $n + 1$. Arguing similar to the proof of Lemma 4.2 and using (4.6), we derive for any $S \leq s \leq t \leq S + T_0$

$$\begin{aligned} \mathbb{E}^s |\varphi_t^{(n+1)} - \mathbb{E}^s \varphi_t^{(n+1)}| &\leq 2 \mathbb{E}^s \left| \varphi_t^{(n+1)} - \varphi_s^{(n+1)} - \int_s^t b(L_r + \mathbb{E}^s \varphi_r^{(n)}) dr \right| \\ &= 2 \mathbb{E}^s \left| \int_s^t (b(L_r + \varphi_r^{(n)}) - b(L_s + \mathbb{E}^s \varphi_r^{(n)})) dr \right| \\ &\leq 4 \|b\|_{C^\beta} \int_s^t (\mathbb{E}^s [|L_r - L_s|^{\beta \wedge 1} \wedge 1] + (\mathbb{E}^s |\varphi_r^{(n)} - \mathbb{E}^s \varphi_r^{(n)}|^{\beta \wedge 1})) dr \\ &\leq 4 \|b\|_{C^\beta} \left(M^* |t - s|^{1 + (\frac{\beta \wedge 1 \wedge \alpha}{\alpha} - \varepsilon)} + (8 \|b\|_{C^\beta} M^*)^{\beta \wedge 1} |t - s|^{1 + (\beta \wedge 1)m} \right) \end{aligned} \quad (\text{A.10})$$

$$\begin{aligned} &\leq 4 \|b\|_{C^\beta} M^* \left(|t - s|^m + (8 \|b\|_{C^\beta})^{\beta \wedge 1} |t - s|^m T_0^{1 + (\beta \wedge 1)m - m} \right) \quad (\text{A.11}) \\ &\leq 8 \|b\|_{C^\beta} M^* |t - s|^m, \end{aligned}$$

where in (A.10) we used the induction step and assumption H3, in (A.11) we used (A.8), and in the last inequality we used (A.9). This proves (A.7) for $n + 1$, and thus completes the proof. \square

Now we can obtain the existence part of Theorem 2.1.

Proof of Theorem 2.1: strong existence and convergence of the Picard iterations. Fix $p \geq 2$ and $\delta > 0$ small enough such that (A.4) holds.

Step 1. For any $n \in \mathbb{Z}_+$ we apply Lemma A.1 to $\varphi^{(n)}$ and $\varphi^{(n+1)}$ with $\tau = 1 + (\frac{\beta \wedge \alpha \wedge 1}{\alpha}) - \delta$; as in the proof of uniqueness part it is clear that $\tau > (2 - \beta)/\alpha$. We see that by Lemma A.2 condition (A.1) holds for some T_0 not depending on n , $C_\varphi = C_\psi = N$, where N also does not depend on n . Thus, all the conditions of Lemma A.1 are met and we have for any $n \in \mathbb{Z}_+$, $T \in (0, T_0]$, $S \in [0, 1 - T]$,

$$\begin{aligned} \|\varphi^{(n+2)} - \varphi^{(n+1)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [S, T]} &= \|\mathcal{S}\varphi^{(n+1)} - \mathcal{S}\varphi^{(n)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [S, T]} \\ &\leq NT^\varepsilon \|\varphi^{(n+1)} - \varphi^{(n)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [S, T]} + \|\varphi^{(n+2)}(S) - \varphi^{(n+1)}(S)\|_{L_p(\Omega)}, \end{aligned}$$

where $N = N(\alpha, \beta, p, d, \|b\|_{C^\beta})$, $\varepsilon = \varepsilon(\alpha, \beta)$.

Pick now $T \in (0, T_0]$ small enough so that $NT^\varepsilon \leq 1/2$. We stress that the choice of T does not depend on n . Let $M := \lceil \frac{1}{T} \rceil$ and let $0 = S_0 \leq S_1 \leq \dots \leq S_M = 1$ be a partition of $[0, 1]$ such that $S_{m+1} - S_m \leq T$, $m = 1, \dots, M$. Then for any $m = 0, \dots, M - 1$ we get

$$\begin{aligned} \|\varphi^{(n+2)} - \varphi^{(n+1)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [S_m, S_{m+1}]} &\leq \frac{1}{2} \|\varphi^{(n+1)} - \varphi^{(n)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [S_m, S_{m+1}]} \\ &\quad + \|\varphi^{(n+2)}(S_m) - \varphi^{(n+1)}(S_m)\|_{L_p(\Omega)}. \end{aligned} \quad (\text{A.12})$$

Step 2. We claim that for any $m = 0, \dots, M - 1$,

$$\|\varphi^{(n+1)} - \varphi^{(n)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [S_m, S_{m+1}]} \leq (m + 1)n^m 2^{-n} \|\varphi^{(1)} - \varphi^{(0)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [0, 1]}, \quad n \in \mathbb{Z}_+. \quad (\text{A.13})$$

We show (A.13) by induction over m . If $m = 0$ then $S_m = 0$ and $\varphi^{(n+1)}(0) = \varphi^{(n)}(0) = \eta$ for any $n \in \mathbb{Z}_+$. Hence, we immediately get from (A.12)

$$\|\varphi^{(n+1)} - \varphi^{(n)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [S_0, S_1]} \leq 2^{-n} \|\varphi^{(1)} - \varphi^{(0)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [S_0, S_1]}, \quad n \in \mathbb{Z}_+,$$

which yields (A.13).

Assume now that (A.13) holds for some $m - 1 \in \{0, 1, \dots, M - 2\}$. We will show that (A.13) holds for m . We have from the induction assumption for any $n \in \mathbb{Z}_+$

$$\begin{aligned} \|\varphi^{(n+1)}(S_m) - \varphi^{(n)}(S_m)\|_{L_p(\Omega)} &\leq \|\varphi^{(n+1)}(S_m) - \varphi^{(n)}(S_m)\|_{\mathcal{C}_p^{\frac{1}{2}}, [S_{m-1}, S_m]} \\ &\leq mn^{m-1} 2^{-n} \|\varphi^{(1)} - \varphi^{(0)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [0, 1]}. \end{aligned}$$

Therefore, (A.12) yields for any $n \in \mathbb{N}$

$$\|\varphi^{(n+1)} - \varphi^{(n)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [S_m, S_{m+1}]} \leq \frac{1}{2} \|\varphi^{(n)} - \varphi^{(n-1)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [S_m, S_{m+1}]} + mn^{m-1} 2^{-n} \|\varphi^{(1)} - \varphi^{(0)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [0, 1]}.$$

Iterating this inequality over n (m is fixed!)

$$\begin{aligned} \|\varphi^{(n+1)} - \varphi^{(n)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [S_m, S_{m+1}]} &\leq 2^{-n} \|\varphi^{(1)} - \varphi^{(0)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [0, 1]} \\ &\quad + m \sum_{i=1}^n 2^{-n+i} 2^{-i} i^{m-1} \|\varphi^{(1)} - \varphi^{(0)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [0, 1]} \\ &= 2^{-n} \|\varphi^{(1)} - \varphi^{(0)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [0, 1]} (1 + m \sum_{i=1}^n i^{m-1}). \end{aligned} \quad (\text{A.14})$$

Clearly,

$$1 + m \sum_{i=1}^n i^{m-1} \leq 1 + mn^m \leq (m+1)n^m,$$

which together with (A.14) gives

$$\|\varphi^{(n+1)} - \varphi^{(n)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [S_m, S_{m+1}]} \leq 2^{-n} \|\varphi^{(1)} - \varphi^{(0)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [0, 1]} (m+1)n^m,$$

which is the desired claim (A.13).

Step 3. Note that by definition of $\varphi^{(1)}$ and $\varphi^{(0)}$

$$\|\varphi^{(1)} - \varphi^{(0)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [0, 1]} \leq \|b\|_{\mathcal{C}^0}.$$

Therefore, using the bound $(m+1)n^m 2^{-n} \leq N 2^{-\frac{n}{2}}$ whenever $m \leq M$ for some $N = N(M) = N(\alpha, \beta, p, d, \|b\|_{\mathcal{C}^\beta})$, we rewrite (A.13) as

$$\|\varphi^{(n+1)} - \varphi^{(n)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [S_m, S_{m+1}]} \leq N 2^{-\frac{n}{2}}, \quad n \in \mathbb{Z}_+,$$

which yields

$$\|\varphi^{(n+1)} - \varphi^{(n)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [0, 1]} \leq N 2^{-\frac{n}{2}}, \quad n \in \mathbb{Z}_+,$$

and eventually

$$\|\varphi^{(k)} - \varphi^{(n)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [0,1]} \leq N2^{-\frac{n \wedge k}{2}}, \quad n, k \in \mathbb{Z}_+. \quad (\text{A.15})$$

Step 4. It follows by (A.15)

$$\sup_{t \in [0,1]} \|\varphi_t^{(k)} - \varphi_t^{(n)}\|_{L_p(\Omega)} \leq \|\varphi^{(k)} - \varphi^{(n)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [0,1]} \leq N2^{-\frac{n \wedge k}{2}}, \quad n, k \in \mathbb{Z}_+. \quad (\text{A.16})$$

Thus, for any $t \in [0, 1]$ the sequence $(\varphi_t^{(n)} - \eta)_{n \in \mathbb{Z}_+} = (\varphi_t^{(n)} - \varphi_t^{(0)})_{n \in \mathbb{Z}_+}$ converges in $L_p(\Omega)$ as $n \rightarrow \infty$. Denote its limit by ψ_t . We have from (A.15) by Fatou's lemma

$$\sup_{t \in [0,1]} \|\psi_t - (\varphi_t^{(n)} - \eta)\|_{L_p(\Omega)} \leq \|\psi - (\varphi^{(n)} - \eta)\|_{\mathcal{C}_p^{\frac{1}{2}}, [0,1]} \leq N2^{-\frac{n}{2}}. \quad (\text{A.17})$$

Put $X_t := \eta + \psi_t + L_t$. We claim that a version of X is a strong solution to (1.1) on $[0, 1]$. Indeed, for any $t \in [0, 1]$ random vector X_t is clearly \mathcal{F}_t -measurable as a limit of \mathcal{F}_t -measurable random vectors; further, recalling (A.6) we have

$$\begin{aligned} \left\| X_t - \eta - \int_0^t b(X_s) ds - L_t \right\|_{L_2(\Omega)} &= \left\| \psi_t - \int_0^t b(L_s + \eta + \psi_s) ds \right\|_{L_2(\Omega)} \\ &\leq \left\| \psi_t - (\varphi_t^{(n+1)} - \eta) \right\|_{L_2(\Omega)} \\ &\quad + \left\| \int_0^t (b(L_s + \eta + \psi_s) - b(L_s + \varphi_s^{(n)})) ds \right\|_{L_2(\Omega)} \\ &\leq \left\| \psi_t - (\varphi_t^{(n+1)} - \eta) \right\|_{L_2(\Omega)} + \int_0^t \|\psi_s - (\varphi_s^{(n)} - \eta)\|_{L_2(\Omega)}^{\beta \wedge 1} ds. \end{aligned}$$

By (A.17), the right-hand side of the above inequality tends to 0 as $n \rightarrow \infty$. Thus, for any $t \in [0, 1]$ we have $X_t = \eta + \int_0^t b(X_s) ds + L_t$ a.s. In particular, $\tilde{X}_t := \eta + \int_0^t b(X_s) ds + L_t$ coincides with X for a.e. ω, t and therefore satisfies

$$\mathbb{P}\left(\tilde{X}_t = \eta + \int_0^t b(\tilde{X}_s) ds + L_t \text{ for all } t \in [0, 1]\right) = 1.$$

Hence X is a strong solution to SDE (1.1) on $[0, 1]$.

Step 5. We see from the definition of the Picard approximation $Y^{(n)}$ and $\varphi^{(n)}$ that $Y^{(n)} = \varphi^{(n)} + L$. By definition of \tilde{X} , we have $X_t = \tilde{X}_t$ a.s. for any fixed $t \in [0, 1]$. Therefore, we can rewrite (A.17) as

$$\|\tilde{X} - Y^{(n)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [0,1]} = \|X - Y^{(n)}\|_{\mathcal{C}_p^{\frac{1}{2}}, [0,1]} \leq N2^{-\frac{n}{2}}.$$

Since the process $\tilde{X} - Y^{(n)}$ is continuous for almost all ω , the Kolmogorov continuity theorem implies

$$\|\tilde{X} - Y^{(n)}\|_{C^0([0,1])} \|X - Y^{(n)}\|_{L_p(\Omega)} \leq N2^{-\frac{n}{2}}.$$

which yields the desired convergence of $\|\tilde{X} - Y^{(n)}\|_{C^0([0,1])}$ to 0 a.s. and in $L_p(\Omega)$. \square

Appendix B. Proofs of the auxiliary statements of the article

Proof of Proposition 2.8. Recall that we denoted the generating triplet of L by (a, Q, ν) . By [SSW12, Theorem 1.3] with $f(s) = s^\alpha$, assumption (2.9) implies that there exists $t_0 = t_0(\alpha) > 0$ such that the gradient bound (2.1) holds for small enough $t \in (0, t_0]$. If $t \in (t_0, 1]$, then

$$\|\nabla \mathcal{P}_t f\|_{C^0} \leq \|\mathcal{P}_{t-t_0} \nabla \mathcal{P}_{t_0} f\|_{C^0} \leq N t_0^{-1/\alpha} \|f\|_{C^0} \leq \tilde{N} t^{-1/\alpha} \|f\|_{C^0},$$

where $\tilde{N} = N t_0^{-1/\alpha}$, which proves (2.1).

Now we move on to H2. Fix $\varepsilon > 0$. Note that by [KS19a, Lemma A.2], (2.1) implies that for any $\varepsilon > 0$

$$\int_{|y| \leq 1} |y|^{\alpha+\varepsilon} \nu(dy) < \infty.$$

Therefore, by [KS19a, Lemma A.3(i)], $Q = 0$.

If $\alpha \in [1, 2)$, then [KS19a, Theorem 3.2(iii)] implies for any $f \in \mathcal{C}^{\alpha+\varepsilon}$ vanishing at infinity

$$\mathcal{L}f(x) = \langle a, \nabla f(x) \rangle + \int_{|y| \geq 1} (f(x+y) - f(x)) \nu(dy) + \int_{|y| \leq 1} \int_0^1 \langle \nabla f(x+\lambda y) - \nabla f(x), y \rangle d\lambda \nu(dy).$$

Hence

$$\|\mathcal{L}f\|_{C^0} \leq |a| \|f\|_{C^1} + 2 \|f\|_{C^0} \nu(\{|y| \geq 1\}) + \|f\|_{\mathcal{C}^{\alpha+\varepsilon}} \int_{|y| \leq 1} |y|^{\alpha+\varepsilon} \nu(dy) \leq N \|f\|_{\mathcal{C}^{\alpha+\varepsilon}}.$$

Very similarly, $\|\mathcal{L}f\|_{C^1} \leq N \|f\|_{\mathcal{C}^{\alpha+1+\varepsilon}}$.

If $\alpha \in (0, 1)$, then by above $\int_{|y| \leq 1} |y| \nu(dy) < \infty$. Consider now the process $\tilde{L}_t := L_t + \kappa t$, where $\kappa = -a + \int_{|y| \leq 1} y \nu(dy)$. Let $\tilde{\mathcal{L}}$ be its generator. It is immediate to see that

$$\mathbb{E} e^{i\langle \lambda, \tilde{L}_t \rangle} = \exp\left(t \int_{\mathbb{R}^d} (e^{i\langle \lambda, y \rangle} - 1) \nu(dy)\right), \quad \lambda \in \mathbb{R}^d, t \geq 0.$$

Therefore all the conditions of [KS19a, Theorem 3.2(ii)] are satisfied and, thus, for any $f \in \mathcal{C}^{\alpha+\varepsilon}$ vanishing at infinity

$$\tilde{\mathcal{L}}f(x) = \int_{\mathbb{R}^d} (f(x+y) - f(y)) \nu(dy).$$

Hence $\|\tilde{\mathcal{L}}f\|_{C^0} \leq \|f\|_{\mathcal{C}^{\alpha+\varepsilon}} \int_{\mathbb{R}^d} (|y|^{\alpha+\varepsilon} \wedge 2) \nu(dy) \leq N \|f\|_{\mathcal{C}^{\alpha+\varepsilon}}$. The bound on $\|\tilde{\mathcal{L}}f\|_{C^1}$ is established by the same argument. \square

Proofs that Examples 2.10 to 2.18 satisfy H1–H3. We begin with Example 2.10. In this case, $\text{Re } \Phi(\lambda) = c_\alpha \int_{\mathbb{S}} |\langle \lambda, \xi \rangle|^\alpha \mu(d\xi)$, for some $c_\alpha > 0$ and the upper bound in (2.9) is immediate. The lower bound follows from the argument presented in [Pri12, p. 424–425 (after Hypothesis 2)]. By Proposition 2.8, this implies that L (or its shifted version) satisfies H1 and H2.

It is easy to see that $\int_{|y| \geq 1} |y|^p \nu(dy) < \infty$ for any $p \in (0, \alpha)$, which implies [Sat13, Theorem 25.3] that

$$\mathbb{E}|L_1|^p < \infty, \quad p \in (0, \alpha). \quad (\text{B.1})$$

If $\alpha \in (1, 2)$, then by [Sat13, formula (14.15)], $\text{Law}(L_t - \kappa t) = \text{Law}(t^{1/\alpha}(L_1 - \kappa))$ for some $\kappa \in \mathbb{R}^d$. Thus, using (B.1), we get for $p \in (0, \alpha)$, $t \in (0, 1]$

$$\mathbb{E}|L_t|^p \leq N(t^{p/\alpha} \mathbb{E}|L_1|^p + t^p + t^{p/\alpha}) \leq Nt^{p/\alpha},$$

since $\alpha > 1$. This shows H3.

If $\alpha = 1$, then by [Sat13, formula (14.16)], $\text{Law } L_t = \text{Law}(tL_1 + \kappa t \log t)$ for some $\kappa \in \mathbb{R}^d$. Thus, by (B.1), for any $p \in (0, 1)$ $\mathbb{E}|L_t|^p \leq Nt^{p(1-\varepsilon)}$, which is H3.

Finally, if $\alpha \in (0, 1)$, then consider the process $\tilde{L}_t := L_t + \kappa t$, for $\kappa = \int_{|y| \leq 1} |y| \nu(dy)$ (this integral is finite by [Sat13, Proposition 14.5]). By above, \tilde{L} satisfies H1 and H2. Further, an easy direct calculation (see also [Sat13, Remark 14.6]) shows that \tilde{L} has a symbol

$$\tilde{\Phi}(\lambda) = \int_{\mathbb{S}} \int_0^\infty (1 - e^{i\langle \lambda, r\xi \rangle}) r^{-1-\alpha} dr \mu(d\xi).$$

Hence $\text{Law}(\tilde{L}_t) = t^{1/\alpha} \tilde{L}_t$, and thus thanks to (B.1), $\mathbb{E}|\tilde{L}_t|^p \leq Nt^{p/\alpha}$ for $p \in (0, \alpha)$. Thus, \tilde{L} satisfies H3.

Example 2.11 is a special case of Example 2.10 with μ being uniform measure on \mathbb{S} . Similarly, Example 2.12 is a special case of Example 2.10 with $\mu = \sum_{k=1}^d (\delta_{e_k} + \delta_{-e_k})$, where (e_k) is the standard basis in \mathbb{R}^d , see, e.g., [Pri12, p. 425].

Now let us move on to Example 2.13. The Lévy process L now has the symbol

$$\Phi(\lambda) = \int_0^\infty \int_{\mathbb{S}} (1 - \cos(r\langle \lambda, \xi \rangle)) r^{-1-\alpha} \rho(r) \mu(d\xi) dr,$$

where we used the fact that μ is symmetric. It is easy to see that non-degeneracy of μ implies that $\int_{\mathbb{S}} |\langle \lambda, \xi \rangle| \mu(d\xi) > 0$ for any $\lambda \in \mathbb{S}$ and thus

$$\inf_{\lambda \in \mathbb{S}} \int_{\mathbb{S}} |\langle \lambda, \xi \rangle| \mu(d\xi) > 0. \quad (\text{B.2})$$

Applying (2.12), we get for any $\lambda \in \mathbb{R}^d$

$$\Phi(\lambda) \leq N \int_{\mathbb{S}} \int_0^\infty (1 - \cos(r\langle \lambda, \xi \rangle)) r^{-1-\alpha} dr \mu(d\xi) \leq N|\lambda|^\alpha \int_{\mathbb{S}} |\langle \frac{\lambda}{|\lambda|}, \xi \rangle|^\alpha \mu(d\xi) \leq N|\lambda|^\alpha. \quad (\text{B.3})$$

Similarly, denoting $\bar{\lambda} := \frac{\lambda}{|\lambda|}$, we get for any $\lambda \in \mathbb{R}^d$, with $|\lambda| > 1/C$

$$\begin{aligned} \Phi(\lambda) &\geq \int_{\mathbb{S}} \int_0^C (1 - \cos(r\langle \lambda, \xi \rangle)) r^{-1-\alpha} dr \mu(d\xi) \\ &= |\lambda|^\alpha \int_{\mathbb{S}} |\langle \bar{\lambda}, \xi \rangle|^\alpha \int_0^{C|\langle \bar{\lambda}, \xi \rangle|} (1 - \cos r) r^{-1-\alpha} dr \mu(d\xi) \\ &\geq \frac{1}{10} |\lambda|^\alpha \int_{\mathbb{S}} |\langle \bar{\lambda}, \xi \rangle|^\alpha \int_0^{|\langle \bar{\lambda}, \xi \rangle|} r^{1-\alpha} dr \mu(d\xi) \\ &\geq N|\lambda|^\alpha \int_{\mathbb{S}} |\langle \bar{\lambda}, \xi \rangle|^2 \mu(d\xi) \\ &\geq N|\lambda|^\alpha \left(\int_{\mathbb{S}} |\langle \bar{\lambda}, \xi \rangle| \mu(d\xi) \right)^2 \end{aligned}$$

$$\geq N|\lambda|^\alpha.$$

where we used inequality $1 - \cos r > \frac{r^2}{10}$ valid for $r \in [0, 1]$ and (B.2). Combining this with (B.3), we see that symbol Φ satisfies (2.9) for large $|\lambda|$ and (2.10) for all $\lambda \in \mathbb{R}^d$. Thus, L satisfies H1–H3.

Examples 2.14 and 2.15 are special cases of Example 2.13.

In case of Example 2.16, we have $N^{-1}|\lambda|^\alpha \leq \Phi(\lambda) \leq N|\lambda|^\alpha$ for large enough $|\lambda|$. Further, $\Phi(\lambda) \leq N\lambda^2$ for small enough $|\lambda|$. Thus, by Proposition 2.8 and Proposition 2.9, assumptions H1–H3 holds for the relativistic α -stable process with $\tilde{\alpha} = \alpha$.

Example 2.17, namely when L is a Brownian motion, is obvious.

Finally, let us treat Example 2.18. We begin with part (i). Without loss of generality, assume $\alpha_1 \leq \alpha_2$. Put $L := L^{(1)} + L^{(2)}$. Let \mathcal{P} (respectively $\mathcal{P}^{(i)}$) be the semigroup associated with L (respectively $L^{(i)}$, $i = 1, 2$). Similarly let \mathcal{L} , $\mathcal{L}^{(i)}$ be the generators of L , respectively $L^{(i)}$. Then for any $f \in \mathcal{C}^0(\mathbb{R}^d)$, $x \in \mathbb{R}^d$ we have by independence of $L^{(1)}$ and $L^{(2)}$

$$\mathcal{P}_t f(x) = \mathbb{E} \mathcal{P}_t^{(2)} f(x + L_t^{(1)}).$$

Thus, since $L^{(2)}$ satisfies H1,

$$|\nabla \mathcal{P}_t f(x)| = |\mathbb{E} \nabla \mathcal{P}_t^{(2)} f(x + L_t^{(1)})| \leq N t^{-1/\alpha_2} \|f\|_{\mathcal{C}^0}.$$

Therefore, L satisfies H1 with $\alpha = \alpha_2$. Since $\mathcal{L} = \mathcal{L}^{(1)} + \mathcal{L}^{(2)}$, we have for any $\varepsilon > 0$, $f \in \mathcal{C}^{\alpha_2 + \delta + \varepsilon}(\mathbb{R}^d)$ vanishing at infinity, $\delta = 0, 1$

$$\|\mathcal{L} f\|_{\mathcal{C}^\delta(\mathbb{R}^d)} \leq \|\mathcal{L}^{(1)} f\|_{\mathcal{C}^\delta(\mathbb{R}^d)} + \|\mathcal{L}^{(2)} f\|_{\mathcal{C}^\delta(\mathbb{R}^d)} \leq M \|f\|_{\mathcal{C}^{\alpha_1 + \delta + \varepsilon}} + M \|f\|_{\mathcal{C}^{\alpha_2 + \delta + \varepsilon}} \leq 2M \|f\|_{\mathcal{C}^{\alpha_2 + \delta + \varepsilon}},$$

and thus L satisfies H2 with $\alpha = \alpha_2$. Finally, to verify H3 we fix $p \in (0, \alpha_2)$. Then If $p < \alpha_1$, then there is nothing to prove:

$$\mathbb{E}[|L_t^{(1)} + L_t^{(2)}|^p \wedge 1] \leq N \mathbb{E}|L_t^{(1)}|^p + N \mathbb{E}|L_t^{(2)}|^p \leq N t^{p/\alpha_2 - \varepsilon},$$

since $\alpha_2 \geq \alpha_1$. Alternatively, if $\alpha_1 \leq p < \alpha_2$, then

$$\mathbb{E}[|L_t^{(1)} + L_t^{(2)}|^p \wedge 1] \leq N \mathbb{E}|L_t^{(1)}|^{\alpha_1 - \varepsilon \alpha_1} + N \mathbb{E}|L_t^{(2)}|^p \leq N(t^{1-2\varepsilon} + t^{\frac{p}{\alpha_2} - \varepsilon}) \leq N t^{\frac{p}{\alpha_2} - 2\varepsilon},$$

where in the last inequality we used that $\frac{p}{\alpha_2} < 1$.

Part (ii) of Example 2.18 is immediate.

To treat part (iii) of Example 2.18, we recall that by [App09, Theorem 2.4.16] Lévy–Itô decomposition holds. Namely, there exists a d -dimensional Brownian motion W with covariance matrix Q and an independent Poisson random measure N such that

$$L_t = W_t + J_t = W_t + at + \int_0^t \int_{|x| \geq 1} x N(dr, dx) + \int_0^t \int_{|x| \leq 1} x \tilde{N}(dr, dx), \quad (\text{B.4})$$

where \tilde{N} is the compensated Poisson measure: $\tilde{N}(D_1, D_2) := N(D_1, D_2) - \text{Leb}(D_1)\nu(D_2)$, $D_1 \in \mathcal{B}(\mathbb{R}_+)$, $D_2 \in \mathcal{B}(\mathbb{R}^d)$, Leb is the Lebesgue measure on \mathbb{R}_+ .

By noting, that the processes W and J are independent, it is easy to see that H1 is satisfied with $\alpha = 2$ by exactly the same argument as in the proof of part (i) above.

H2 is satisfied with $\alpha = 2$ by the same argument as we used in the proof of Proposition 2.8; recall that $\int_{|y| \leq 1} |y|^2 \nu(dy) < \infty$ thanks to the requirements to the jump measure.

To show H3, we note that for any $p \in (0, 2]$

$$\mathbb{E}|W_t|^p \leq Nt^{p/2}; \quad (\text{B.5})$$

$$\mathbb{E} \left| \int_0^t \int_{|x| \leq 1} x \tilde{N}(dr, dx) \right|^p \leq Nt^{p/2} \left(\int_{|x| \leq 1} x^2 \nu(dx) \right)^{p/2} = Nt^{p/2}, \quad (\text{B.6})$$

where the inequality follows from, e.g., [Kun04, Lemma 2.4]. Further, by definition

$$\begin{aligned} \mathbb{E} \left| \int_0^t \int_{|x| \geq 1} x N(dr, dx) \right|^{\gamma \wedge 1} &= \mathbb{E} \left| \sum_{0 \leq s \leq t} \Delta L_s \mathbb{1}(|\Delta L_s| \geq 1) \right|^{\gamma \wedge 1} \\ &\leq \mathbb{E} \sum_{0 \leq s \leq t} |\Delta L_s|^{\gamma \wedge 1} \mathbb{1}(|\Delta L_s| \geq 1) \\ &= t \int_{|x| \geq 1} |x|^{\gamma \wedge 1} \nu(dx) = Nt. \end{aligned}$$

If $p \leq (\gamma \wedge 1)$, then combining this with (B.5), (B.6) and substituting into the decomposition (B.4), we get

$$\mathbb{E}|L_t|^p \leq Nt^{p/2} + Nt^{p/(\gamma \wedge 1)} \leq Nt^{p/2}.$$

If $(\gamma \wedge 1) \leq p \leq 2$, then similarly

$$\mathbb{E}[|L_t|^p \wedge 1] \leq Nt^{p/2} + Nt \leq Nt^{p/2}.$$

Thus L satisfies H3 with $\alpha = 2$. □

Proof of Lemma 3.1. The proof is based on the method of [Ger20, proof of Lemma 2.2] (see also the proof of the original stochastic sewing lemma in [Lê20]). The novelty here is that we apply the conditional Burkholder-Davis-Gundy inequality rather than the standard one. Let $(Z_i)_{i=2, \dots, M}$ be a sequence of random vectors in \mathbb{R}^d adapted to the filtration $\mathbb{G} := (\mathcal{G}_i)_{i \geq 0}$, and assume that $\mathcal{G} \subset \mathcal{G}_0$. Then

$$\sum_{i=2}^M Z_i = \sum_{i=2}^M \mathbb{E}^{\mathcal{G}_{i-2}} Z_i + \sum_{i \text{ even}} (Z_i - \mathbb{E}^{\mathcal{G}_{i-2}} Z_i) + \sum_{i \text{ odd}} (Z_i - \mathbb{E}^{\mathcal{G}_{i-2}} Z_i). \quad (\text{B.7})$$

The sequence $(Z_i - \mathbb{E}^{\mathcal{G}_{i-2}} Z_i)_{i \in \{2, \dots, M\}, i \text{ even}}$ is a martingale difference sequence with respect to the filtration $(\mathcal{G}_i)_{i \in \{2, \dots, M\}, i \text{ even}}$. We apply the conditional Burkholder-Davis-Gundy inequality [CF16, Proposition 27]:

$$\begin{aligned} \left\| \sum_{i \text{ even}} (Z_i - \mathbb{E}^{\mathcal{G}_{i-2}} Z_i) \right\|_{L^p(\Omega)|\mathcal{G}} &\leq N \left\| \sum_{i \text{ even}} (Z_i - \mathbb{E}^{\mathcal{G}_{i-2}} Z_i)^2 \right\|_{L^{p/2}(\Omega)|\mathcal{G}}^{\frac{1}{2}} \\ &\leq N \left(\sum_{i \text{ even}} (\|Z_i\|_{L^p(\Omega)|\mathcal{G}}^2 + \|\mathbb{E}^{\mathcal{G}_{i-2}} Z_i\|_{L^p(\Omega)|\mathcal{G}}^2) \right)^{\frac{1}{2}} \\ &\leq N \left(\sum_{i \text{ even}} \|Z_i\|_{L^p(\Omega)|\mathcal{G}}^2 \right)^{\frac{1}{2}}, \end{aligned}$$

where in the last step we used (3.2) and the fact that $\mathcal{G} \subset \mathcal{G}_{i-2}$. Treating the term in (B.7) with the odd terms in the same way, we finally get

$$\left\| \sum_{i=2}^M Z_i \right\|_{L_p(\Omega)|\mathcal{G}} \leq \sum_{i=2}^M \|\mathbb{E}^{\mathcal{G}_{i-2}} Z_i\|_{L_p(\Omega)|\mathcal{G}} + N \left(\sum_{i=2}^M \|Z_i\|_{L_p(\Omega)|\mathcal{G}}^2 \right)^{1/2}. \quad (\text{B.8})$$

Now we proceed to the proof. For $0 \leq s \leq t \leq 1$, $m \in \mathbb{N}$, consider the uniform partition of $[s, t]$: $t_i^m := s + i(t-s)2^{-m}$, $i = 0, \dots, 2^m$. Put

$$A_{s,t}^m := \sum_{i=1}^{2^m-1} A_{t_i^m, t_{i+1}^m}^m, \quad m \in \mathbb{N}.$$

Then it follows from (3.6) and the conditional Fatou's lemma that

$$\|\mathcal{A}_t - \mathcal{A}_s\|_{L_p(\Omega)|\mathcal{G}} \leq \liminf_{m \rightarrow \infty} \|A_{s,t}^m\|_{L_p(\Omega)|\mathcal{G}}. \quad (\text{B.9})$$

For $m \in \mathbb{N}$ we apply (B.8) with $M := 2^m$, $Z_i := \delta A_{t_{i-1}^m, t_{2i-1}^{m+1}, t_i^m}$, $\mathcal{G}_i := \mathcal{F}_{t_i^m}$. We get

$$\begin{aligned} \|A_{s,t}^{m+1} - A_{s,t}^m\|_{L_p(\Omega)|\mathcal{G}} &\leq \left\| \sum_{i=2}^{2^m} \delta A_{t_{i-1}^m, t_{2i-1}^{m+1}, t_i^m} \right\|_{L_p(\Omega)|\mathcal{G}} + \|A_{t_1^{m+1}, t_2^{m+1}}\|_{L_p(\Omega)|\mathcal{G}} \\ &\leq N \sum_{i=2}^{2^m} \|\mathbb{E}^{\mathcal{F}_{t_{i-2}^m}} \delta A_{t_{i-1}^m, t_{2i-1}^{m+1}, t_i^m}\|_{L_p(\Omega)|\mathcal{G}} \\ &\quad + N \left(\sum_{i=2}^{2^m} \|\delta A_{t_{i-1}^m, t_{2i-1}^{m+1}, t_i^m}\|_{L_p(\Omega)|\mathcal{G}}^2 \right)^{1/2} + \|A_{t_1^{m+1}, t_2^{m+1}}\|_{L_p(\Omega)|\mathcal{G}} \\ &\leq N 2^m \Gamma_2 2^{-m(1+\varepsilon_2)} |t-s|^{1+\varepsilon_2} + N 2^m \Gamma_3 2^{-m(1+\varepsilon_3)} |t-s|^{1+\varepsilon_3} \\ &\quad + N 2^{m/2} \Gamma_1 2^{-m(1/2+\varepsilon_1)} |t-s|^{1/2+\varepsilon_1} + \Gamma_1 2^{-m(1/2+\varepsilon_1)} |t-s|^{1/2+\varepsilon_1} \\ &\leq N \Gamma_2 2^{-m\varepsilon_2} |t-s|^{1+\varepsilon_2} + N \Gamma_3 2^{-m\varepsilon_3} |t-s|^{1+\varepsilon_3} \\ &\quad + N \Gamma_1 2^{-m\varepsilon_1} |t-s|^{1/2+\varepsilon_1} + \Gamma_1 2^{-m(1/2+\varepsilon_1)} |t-s|^{1/2+\varepsilon_1}, \quad (\text{B.10}) \end{aligned}$$

where in the penultimate inequality we used bounds (3.4) and (3.5), the fact that $t_{i-1}^m - (t_i^m - t_{i-1}^m) = t_{i-2}^m$ and inequality

$$\|\delta A_{s,u,t}\|_{L_p(\Omega)|\mathcal{G}} \leq \|A_{s,t}\|_{L_p(\Omega)|\mathcal{G}} + \|A_{s,u}\|_{L_p(\Omega)|\mathcal{G}} + \|A_{u,t}\|_{L_p(\Omega)|\mathcal{G}}.$$

Summing up inequalities (B.10) over m and using (3.4) once again, we deduce for $m \in \mathbb{N}$

$$\begin{aligned} \|A_{s,t}^m\|_{L_p(\Omega)|\mathcal{G}} &\leq \sum_{i=1}^{m-1} \|A_{s,t}^{i+1} - A_{s,t}^i\|_{L_p(\Omega)|\mathcal{G}} + \|A_{s,t}^1\|_{L_p(\Omega)|\mathcal{G}} \\ &\leq N \Gamma_2 |t-s|^{1+\varepsilon_2} + N \Gamma_3 |t-s|^{1+\varepsilon_3} + N \Gamma_1 |t-s|^{1/2+\varepsilon_1}. \end{aligned}$$

This together with (B.9) yields the desired estimate (3.7). \square

The proof of Proposition 3.2 relies on the following inequality which is a very minor modification of [Lê22a, Lemma 2.1].

Proposition B.1. *Let X, Y be nonnegative random variables. Assume that for any $\alpha > 0, \theta \in (0, 1)$ one has*

$$\mathbb{P}(X \geq 2\alpha) \leq \theta \mathbb{P}(X \geq \alpha) + \mathbb{P}(Y \geq \theta\alpha). \quad (\text{B.11})$$

Then for any $p \geq 1$ there exists a constant $N = N(p)$ such that

$$\|X\|_{L_p(\Omega)} \leq N(p)\|Y\|_{L_p(\Omega)}. \quad (\text{B.12})$$

Proof. We follow the proof of [Lê22a, Lemma 2.1]. Using the identity $x^p = p \int_0^x \lambda^{p-1} d\lambda$ valid for any $x \geq 0, p \geq 1$, we derive for any $\theta \in (0, 1), M > 0$

$$\begin{aligned} \mathbb{E}(X \wedge M)^p &= p \int_0^M \mathbb{P}(X \geq \lambda) \lambda^{p-1} d\lambda = 2^p p \int_0^{\frac{M}{2}} \mathbb{P}(X \geq 2\lambda) \lambda^{p-1} d\lambda \\ &\leq 2^p p \theta \int_0^{\frac{M}{2}} \mathbb{P}(X \geq \lambda) \lambda^{p-1} d\lambda + 2^p p \int_0^{\frac{M}{2}} \mathbb{P}(Y \geq \theta\lambda) \lambda^{p-1} d\lambda \\ &\leq 2^p \theta \mathbb{E}(X \wedge M)^p + (2\theta^{-1})^p \mathbb{E}Y^p, \end{aligned}$$

where in the penultimate inequality we use our main assumption (B.11). Take now $\theta := 2^{-p-1}$. Then, by above

$$\frac{1}{2} \mathbb{E}(X \wedge M)^p \leq 2^{p^2+2p} \mathbb{E}Y^p.$$

Passing to the limit as $M \rightarrow \infty$ and using the monotone convergence theorem, we obtain (B.12). \square

Proof of Proposition 3.2. Fix $0 \leq S \leq T$. Put

$$\xi^*(\omega) := \sup_{r \in [S, T]} \xi_r(\omega); \quad V^*(\omega) := \sup_{r \in [S, T]} |\mathcal{A}_r(\omega) - A_S(\omega)|, \quad \omega \in \Omega.$$

We will assume that ξ is bounded from below by a deterministic constant $\delta > 0$. This is no loss of generality: if this is not the case, we can replace ξ with $\xi' := \xi \vee \delta$, note that the condition (3.8) still holds with ξ' , apply the claim with ξ' , and pass to the $\delta \rightarrow 0$ limit in the final bound (3.9), which is possible thanks to the monotone convergence theorem.

Step 1. We claim that for any stopping times $\tau \leq \eta$ taking values in $[S, T]$ we have a.s.

$$\mathbb{E}^\tau \frac{|\mathcal{A}_\eta - \mathcal{A}_\tau|}{\xi^*} \leq 2. \quad (\text{B.13})$$

First, we prove (B.13) for the case when $\eta = T$ and τ takes finitely many values $S \leq t_1 < \dots < t_n = T$. Recall the identity $\mathbb{E}^\tau(X \mathbb{1}_{\tau=t}) = \mathbb{E}^t(X \mathbb{1}_{\tau=t})$ valid for any integrable random variable X , see, e.g., [KS91, Problem 1.2.17(i)]. Then, using (3.8) and the fact that $\xi^* \geq \xi_\tau$ we deduce

$$\begin{aligned} \mathbb{E}^\tau \frac{|\mathcal{A}_T - \mathcal{A}_\tau|}{\xi^*} &\leq \mathbb{E}^\tau \frac{|\mathcal{A}_T - \mathcal{A}_\tau|}{\xi_\tau} = \sum_{i=1}^n \mathbb{E}^{t_i} \left(\frac{|\mathcal{A}_T - \mathcal{A}_{t_i}|}{\xi_{t_i}} \mathbb{1}(\tau = t_i) \right) \\ &= \sum_{i=1}^n \mathbb{1}(\tau = t_i) \frac{1}{\xi_{t_i}} \mathbb{E}^{t_i} |\mathcal{A}_T - \mathcal{A}_{t_i}| \leq 1. \end{aligned} \quad (\text{B.14})$$

Next, for a general stopping time $\tau \in [S, T]$, we consider a sequence of stopping times τ_n taking finitely many values in $[S, T]$ and converging to τ from above. Then, for any $N > 0$ we have using continuity of \mathcal{A} and (B.14):

$$\mathbb{E}^\tau \frac{|\mathcal{A}_T - \mathcal{A}_\tau|}{\xi^*} \leq \liminf_{n \rightarrow \infty} \mathbb{E}^\tau \frac{|\mathcal{A}_T - \mathcal{A}_{\tau_n}|}{\xi^*} = \liminf_{n \rightarrow \infty} \mathbb{E}^\tau \mathbb{E}^{\tau_n} \frac{|\mathcal{A}_T - \mathcal{A}_{\tau_n}|}{\xi^*} \leq 1. \quad (\text{B.15})$$

Finally, let $\tau \leq \eta$ be arbitrarily stopping times taking values in $[S, T]$. Then (B.15) yields

$$\mathbb{E}^\tau \frac{|\mathcal{A}_\eta - \mathcal{A}_\tau|}{\xi^*} \leq \mathbb{E}^\tau \frac{|\mathcal{A}_T - \mathcal{A}_\tau|}{\xi^*} + \mathbb{E}^\tau \frac{|\mathcal{A}_T - \mathcal{A}_\eta|}{\xi^*} \leq 1 + \mathbb{E}^\tau \mathbb{E}^\eta \frac{|\mathcal{A}_T - \mathcal{A}_\eta|}{\xi^*} \leq 2,$$

which is (B.13).

Step 2. Now we modify the corresponding part of the proof of [Lê22a, Theorem 1.3] to adapt it to our new condition (B.13). For arbitrary $\alpha > 0$ we consider two stopping times:

$$\tau_\alpha := T \wedge \inf\{r \in [S, T] : |\mathcal{A}_r - \mathcal{A}_S| \geq \alpha\}; \quad \eta_\alpha := T \wedge \inf\{r \in [S, T] : |\mathcal{A}_r - \mathcal{A}_S| \geq 2\alpha\}.$$

Then for any $\theta \in (0, 1)$ we clearly have

$$\begin{aligned} \{V^* \geq 2\alpha\} &\subset \{V^* \geq \alpha, |\mathcal{A}_{\eta_\alpha} - \mathcal{A}_{\tau_\alpha}| \geq \alpha\} \\ &\subset \{V^* \geq \alpha, |\mathcal{A}_{\eta_\alpha} - \mathcal{A}_{\tau_\alpha}| \geq 2\theta^{-1}\xi^*\} \cup \{2\theta^{-1}\xi^* \geq \alpha\}. \end{aligned}$$

Note that $\{V^* \geq \alpha\} = \{|\mathcal{A}_{\tau_\alpha} - \mathcal{A}_S| \geq \alpha\}$ and hence the random variable $\mathbb{1}_{\{V^* \geq \alpha\}}$ is $\mathcal{F}_{\tau_\alpha}$ -measurable. Therefore, using also the Chebyshev inequality we get for any \mathcal{F}_S -measurable set G

$$\begin{aligned} \mathbb{P}(V^* \mathbb{1}_G \geq 2\alpha) &= \mathbb{P}(V^* \geq 2\alpha, G) \\ &\leq \mathbb{P}(V^* \geq \alpha, |\mathcal{A}_{\eta_\alpha} - \mathcal{A}_{\tau_\alpha}| \geq 2\theta^{-1}\xi^*, G) + \mathbb{P}(2\xi^* \geq \alpha\theta, G) \\ &\leq \mathbb{E} \mathbb{1}_{V^* \geq \alpha} \mathbb{1}_G \mathbb{E}^{\tau_\alpha} \mathbb{1}_{\{|\mathcal{A}_{\eta_\alpha} - \mathcal{A}_{\tau_\alpha}| \geq 2\theta^{-1}\xi^*\}} + \mathbb{P}(2\xi^* \mathbb{1}_G \geq \alpha\theta) \\ &\leq \theta \mathbb{E} \mathbb{1}_{V^* \geq \alpha} \mathbb{1}_G \mathbb{E}^{\tau_\alpha} \frac{|\mathcal{A}_{\eta_\alpha} - \mathcal{A}_{\tau_\alpha}|}{2\xi^*} + \mathbb{P}(2\xi^* \mathbb{1}_G \geq \alpha\theta) \\ &\leq \theta \mathbb{P}(V^* \mathbb{1}_G \geq \alpha) + \mathbb{P}(2\xi^* \mathbb{1}_G \geq \alpha\theta), \end{aligned}$$

where in the second inequality we used that $\mathcal{F}_S \subset \mathcal{F}_{\tau_\alpha}$ and in the last inequality we used (B.13). Since $\alpha > 0$ and $\theta \in (0, 1)$ were arbitrary, we see that condition (B.11) of Proposition B.1 holds. Hence, (B.12) yields

$$\|V^* \mathbb{1}_G\|_{L_p(\Omega)} \leq N \|\xi^* \mathbb{1}_G\|_{L_p(\Omega)},$$

for $N = N(p)$ independent of G . Since G was an arbitrary \mathcal{F}_S -measurable set, we get

$$\|V^*\|_{L_p(\Omega)|\mathcal{F}_S} \leq N \|\xi^*\|_{L_p(\Omega)|\mathcal{F}_S},$$

which is the desired bound (3.9). \square

Proof of Proposition 3.3. We begin with the proof of (3.11). We proceed by induction. The case $\rho = 0$ is obvious. Indeed $\|\mathcal{P}_t f\|_{\mathcal{C}^0} \leq \|f\|_{\mathcal{C}^0} \leq \|f\|_{\mathcal{C}^\beta}$.

Assume now that the statement is proved for $\rho \in [0, M]$, $M \in \mathbb{Z}_+$. Let us prove it for $\rho \in (M, M+1]$.

Let $\rho = M+1$. If $\beta \geq 1$, then by the definition of the norm

$$\begin{aligned} \|\mathcal{P}_t f\|_{\mathcal{C}^{M+1}} &\leq \|\mathcal{P}_t f\|_{\mathcal{C}^0} + \sup_{i \in \{1, \dots, d\}} \|\partial_i \mathcal{P}_t f\|_{\mathcal{C}^M} \\ &\leq \|f\|_{\mathcal{C}^0} + \sup_{i \in \{1, \dots, d\}} \|\mathcal{P}_t \partial_i f\|_{\mathcal{C}^M} \\ &\leq \|f\|_{\mathcal{C}^0} + N t^{(\beta-M-1) \wedge 0/\alpha} \sup_{i \in \{1, \dots, d\}} \|\partial_i f\|_{\mathcal{C}^{\beta-1}} \\ &\leq N t^{(\beta-M-1) \wedge 0/\alpha} \|f\|_{\mathcal{C}^\beta}, \end{aligned} \tag{B.16}$$

where in the third inequality we have used the induction step. If $\beta = 0$, then

$$\begin{aligned} \|\mathcal{P}_t f\|_{\mathcal{C}^{M+1}} &\leq \|\mathcal{P}_t f\|_{\mathcal{C}^0} + \sup_{i \in \{1, \dots, d\}} \|\partial_i \mathcal{P}_t f\|_{\mathcal{C}^M} \\ &\leq \|f\|_{\mathcal{C}^0} + \sup_{i \in \{1, \dots, d\}} \|\mathcal{P}_{t/2} \partial_i \mathcal{P}_{t/2} f\|_{\mathcal{C}^M} \\ &\leq \|f\|_{\mathcal{C}^0} + N t^{-M/\alpha} \sup_{i \in \{1, \dots, d\}} \|\partial_i \mathcal{P}_{t/2} f\|_{\mathcal{C}^0} \\ &\leq \|f\|_{\mathcal{C}^0} + N t^{-M/\alpha} t^{-1/\alpha} \|f\|_{\mathcal{C}^0} \\ &\leq N t^{-(M+1)/\alpha} \|f\|_{\mathcal{C}^0}, \end{aligned} \tag{B.17}$$

where in the third inequality we used the induction step and in the fourth inequality we used (2.1). Finally, if $\beta \in (0, 1)$, $f \in \mathcal{C}^\beta$, then define for $\lambda > 0$ the interpolation function

$$K(\lambda, f) := \inf_{\substack{f=a+b \\ a \in \mathcal{C}^0(\mathbb{R}^d), b \in \mathcal{C}^1(\mathbb{R}^d)}} (\|a\|_{\mathcal{C}^0} + \lambda \|b\|_{\mathcal{C}^1}).$$

It is well-known, that if $f \in \mathcal{C}^\beta$, then

$$N^{-1} \|f\|_{\mathcal{C}^\beta} \leq \sup_{\lambda > 0} \frac{K(\lambda, f)}{\lambda^\beta} \leq N \|f\|_{\mathcal{C}^\beta}, \tag{B.18}$$

see, e.g., [Lun18, Example 1.8]. Then for any $a \in \mathcal{C}^0(\mathbb{R}^d)$, $b \in \mathcal{C}^1(\mathbb{R}^d)$ such that $f = a + b$ we get using (B.16) and (B.17)

$$\begin{aligned} \|\mathcal{P}_t f\|_{\mathcal{C}^{M+1}} &= \|\mathcal{P}_t(a+b)\|_{\mathcal{C}^{M+1}} \leq N t^{-(M+1)/\alpha} \|a\|_{\mathcal{C}^0} + N t^{-M/\alpha} \|b\|_{\mathcal{C}^1} \\ &\leq N t^{-(M+1)/\alpha} (\|a\|_{\mathcal{C}^0} + t^{1/\alpha} \|b\|_{\mathcal{C}^1}). \end{aligned}$$

Taking infimum over all $a \in \mathcal{C}^0(\mathbb{R}^d)$, $b \in \mathcal{C}^1(\mathbb{R}^d)$ such that $f = a + b$ and using (B.18), we get

$$\|\mathcal{P}_t f\|_{\mathcal{C}^{M+1}} \leq N t^{-(M+1)/\alpha} K(t^{1/\alpha}, f) \leq N \|f\|_{\mathcal{C}^\beta} t^{-(M+1)/\alpha} t^{\beta/\alpha},$$

which is (3.11). Thus, the case $\rho = M+1$, $\beta \geq 0$ is proven.

Finally if $\rho \in (M, M+1)$, then by above and the standard interpolation inequality

$$\|\mathcal{P}_t f\|_{\mathcal{C}^\rho} \leq \|\mathcal{P}_t f\|_{\mathcal{C}^M}^{M+1-\rho} \|\mathcal{P}_t f\|_{\mathcal{C}^{M+1}}^{\rho-M} \leq N \|f\|_{\mathcal{C}^\beta} t^{\frac{(\beta-M) \wedge 0}{\alpha} (M+1-\rho)} t^{\frac{(\beta-M-1) \wedge 0}{\alpha} (\rho-M)}$$

$$\leq N\|f\|_{\mathcal{C}^\beta} t^{\frac{(\beta-\rho)\wedge 0}{\alpha}}.$$

Now let us prove (3.12). We start by assuming additionally that f vanishes at infinity. First note that it follows from assumptions of the proposition that $1 + \varepsilon \geq \mu \geq (\beta - \rho)/\alpha$. Thus

$$\beta - \rho - \alpha - \varepsilon\alpha \leq 0. \quad (\text{B.19})$$

We see that for any $0 \leq s < t$, $x \in \mathbb{R}^d$,

$$\mathcal{P}_t f(x) - \mathcal{P}_s f(x) = \int_s^t \partial_r \mathcal{P}_r f(x) dr = \int_s^t \mathcal{L} \mathcal{P}_r f(x) dr,$$

where we used the fact that for $r > 0$ we have $\mathcal{P}_r f \in \mathcal{C}^2$ by (3.11) and we applied Itô's formula for Lévy processes, see, e.g., [App09, Theorem 4.4.7]. Therefore, by (2.2) and (3.11), taking also into account (B.19) and that $\mathcal{P}_r f$ vanishes at infinity for any $r \geq 0$, we deduce

$$\begin{aligned} \|\mathcal{P}_t f - \mathcal{P}_s f\|_{\mathcal{C}^\rho} &\leq \int_s^t \|\mathcal{L} \mathcal{P}_r f\|_{\mathcal{C}^\rho} dr \leq N \int_s^t \|\mathcal{P}_r f\|_{\mathcal{C}^{\rho+\alpha+\varepsilon\alpha}} dr \\ &\leq N\|f\|_{\mathcal{C}^\beta} \int_s^t r^{\frac{\beta-\rho-\alpha-\varepsilon\alpha}{\alpha}} dr \\ &\leq N\|f\|_{\mathcal{C}^\beta} \int_s^t s^{\frac{\beta-\rho}{\alpha}-\mu} (r-s)^{\mu-1-\varepsilon} dr \\ &\leq N\|f\|_{\mathcal{C}^\beta} s^{\frac{\beta-\rho}{\alpha}-\mu} (t-s)^{\mu-\varepsilon}, \end{aligned}$$

where in the penultimate inequality we used obvious bounds $r \geq s$ and $r \geq r-s$ (we note that the corresponding exponents are nonpositive since $\frac{\beta-\rho}{\alpha} \leq \mu \leq 1 + \varepsilon$), and in the last bound we used that $\mu > \varepsilon$ and the singularity is integrable. This implies (3.12) for the case when f vanishes at infinity.

In general case, take a smooth function $\chi: \mathbb{R}_+ \rightarrow [0, 1]$ such that $\chi(x) = 1$ for $x \in [0, 1]$ and $\chi(x) = 0$ for $x \geq 2$. Then the function $f_n(x) := f(x)\chi(|x|/n)$ vanishes at infinity and $\|f_n\|_{\mathcal{C}^\beta} \leq N\|f\|_{\mathcal{C}^\beta}$ for some N depending only on the choice of χ . By above,

$$\|\mathcal{P}_t f_n - \mathcal{P}_s f_n\|_{\mathcal{C}^\rho} \leq N\|f_n\|_{\mathcal{C}^\beta} s^{\frac{\beta-\rho}{\alpha}-\mu} (t-s)^{\mu-\varepsilon} \leq N\|f\|_{\mathcal{C}^\beta} s^{\frac{\beta-\rho}{\alpha}-\mu} (t-s)^{\mu-\varepsilon}.$$

By the dominated convergence theorem, $(\mathcal{P}_t - \mathcal{P}_s)f_n \rightarrow (\mathcal{P}_t - \mathcal{P}_s)f$ everywhere. Therefore,

$$\|(\mathcal{P}_t - \mathcal{P}_s)f\|_{\mathcal{C}^\rho} \leq \liminf_{n \rightarrow \infty} \|(\mathcal{P}_t - \mathcal{P}_s)f_n\|_{\mathcal{C}^\rho} \leq N\|f\|_{\mathcal{C}^\beta} s^{\frac{\beta-\rho}{\alpha}-\mu} (t-s)^{\mu-\varepsilon}. \quad \square$$

References

- [AJP22] C. AMORINO, A. JARAMILLO, and M. PODOLSKIJ. Optimal estimation of local time and occupation time measure for an α -stable Lévy process. *arXiv preprint arXiv:2210.07672* (2022).
- [AL22] R. ALTMAYER and R. LE GUÉVEL. Optimal L^2 -approximation of occupation and local times for symmetric stable processes. *Electron. J. Stat.* **16**, no. 1, (2022), 2859–2883. <http://dx.doi.org/10.1214/22-ejs2013>.
- [Alt21] R. ALTMAYER. Approximation of occupation time functionals. *Bernoulli* **27**, no. 4, (2021), 2714–2739. <http://dx.doi.org/10.3150/21-BEJ1328>.

- [App09] D. APPLEBAUM. *Lévy processes and stochastic calculus*, vol. 116 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second ed., 2009. <http://dx.doi.org/10.1017/CB09780511809781>.
- [BDG21a] O. BUTKOVSKY, K. DAREIOTIS, and M. GERENCSÉR. Approximation of SDEs: a stochastic sewing approach. *Probability Theory and Related Fields* (2021). <http://dx.doi.org/10.1007/s00440-021-01080-2>.
- [BDG21b] O. BUTKOVSKY, K. DAREIOTIS, and M. GERENCSÉR. Optimal rate of convergence for approximations of SPDEs with non-regular drift. *arXiv e-prints* arXiv:2110.06148. <http://arxiv.org/abs/2110.06148>.
- [CF16] M. COGHI and F. FLANDOLI. Propagation of chaos for interacting particles subject to environmental noise. *Ann. Appl. Probab.* **26**, no. 3, (2016), 1407–1442. <http://dx.doi.org/10.1214/15-AAP1120>.
- [CMS90] R. CARMONA, W. C. MASTERS, and B. SIMON. Relativistic Schrödinger operators: asymptotic behavior of the eigenfunctions. *J. Funct. Anal.* **91**, no. 1, (1990), 117–142. [http://dx.doi.org/10.1016/0022-1236\(90\)90049-Q](http://dx.doi.org/10.1016/0022-1236(90)90049-Q).
- [CSZ18] Z.-Q. CHEN, R. SONG, and X. ZHANG. Stochastic flows for Lévy processes with Hölder drifts. *Revista Matemática Iberoamericana* **34**, no. 4, (2018), 1755–1788. <http://dx.doi.org/10.4171/rmi/1042>.
- [CVST20] P. CHAKRABORTY, B. VERMEERSCH, A. SHAKOURI, and S. TINDEL. Relativistic stable processes in quasiballistic heat conduction in semiconductors. *Physical Review E* **101**, no. 4, (2020), 042110.
- [DGL21] K. DAREIOTIS, M. GERENCSÉR, and K. LÊ. Quantifying a convergence theorem of Gyöngy and Krylov. *arXiv e-prints* arXiv:2101.12185. <http://arxiv.org/abs/2101.12185>.
- [DS15] C.-S. DENG and R. L. SCHILLING. On shift Harnack inequalities for subordinate semigroups and moment estimates for Lévy processes. *Stochastic Process. Appl.* **125**, no. 10, (2015), 3851–3878. <http://dx.doi.org/10.1016/j.spa.2015.05.013>.
- [FKM21] N. FRIKHA, V. KONAKOV, and S. MENOZZI. Well-posedness of some non-linear stable driven SDEs. *Discrete Contin. Dyn. Syst.* **41**, no. 2, (2021), 849–898. <http://dx.doi.org/10.3934/dcds.2020302>.
- [Ger20] M. GERENCSÉR. Regularisation by regular noise. *arXiv e-prints* (2020). <http://arxiv.org/abs/2009.08418>.
- [GK14] I. GANYCHENKO and A. KULIK. Rates of approximation of nonsmooth integral-type functionals of Markov processes. *Mod. Stoch. Theory Appl.* **1**, no. 2, (2014), 117–126. <http://dx.doi.org/10.15559/vmsta-2014.12>.
- [GKK15] I. GANYCHENKO, V. KNOPOVA, and A. KULIK. Accuracy of discrete approximation for integral functionals of Markov processes. *Mod. Stoch. Theory Appl.* **2**, no. 4, (2015), 401–420. <http://dx.doi.org/10.15559/15-VMSTA46>.
- [HSY21] X. HUANG, Y. SUO, and C. YUAN. Estimate of heat kernel for Euler-Maruyama scheme of SDEs driven by α -stable noise and applications. *arXiv e-prints* arXiv:2103.01323. <http://arxiv.org/abs/2103.01323>.
- [KP02] K. KUBILIUS and E. PLATEN. Rate of weak convergence of the Euler approximation for diffusion processes with jumps. *Monte Carlo Methods and Appl.* **8**, no. 1, (2002), 83–96. <http://dx.doi.org/doi:10.1515/mcma.2002.8.1.83>.
- [KS91] I. KARATZAS and S. E. SHREVE. *Brownian motion and stochastic calculus*, vol. 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second ed., 1991. <http://dx.doi.org/10.1007/978-1-4612-0949-2>.

- [KS19a] F. KÜHN and R. L. SCHILLING. On the domain of fractional Laplacians and related generators of Feller processes. *J. Funct. Anal.* **276**, no. 8, (2019), 2397–2439. <http://dx.doi.org/10.1016/j.jfa.2018.12.011>.
- [KS19b] F. KÜHN and R. L. SCHILLING. Strong convergence of the Euler-Maruyama approximation for a class of Lévy-driven SDEs. *Stochastic Process. Appl.* **129**, no. 8, (2019), 2654–2680. <http://dx.doi.org/10.1016/j.spa.2018.07.018>.
- [Kul19] A. M. KULIK. On weak uniqueness and distributional properties of a solution to an SDE with α -stable noise. *Stochastic Process. Appl.* **129**, no. 2, (2019), 473–506. <http://dx.doi.org/10.1016/j.spa.2018.03.010>.
- [Kun04] H. KUNITA. Stochastic differential equations based on Lévy processes and stochastic flows of diffeomorphisms. In *Real and stochastic analysis*, Trends Math., 305–373. Birkhäuser Boston, Boston, MA, 2004.
- [Lê20] K. LÊ. A stochastic sewing lemma and applications. *Electron. J. Probab.* **25**, (2020), Paper No. 38, 55. <http://dx.doi.org/10.1214/20-ejp442>.
- [Lê22a] K. LÊ. Maximal inequalities and weighted BMO processes. *arXiv preprint arXiv:2211.15550* (2022).
- [Lê22b] K. LÊ. Quantitative john–nirenberg inequality for stochastic processes of bounded mean oscillation. *arXiv preprint arXiv:2210.15736* (2022).
- [LL21] K. LÊ and C. LING. Taming singular stochastic differential equations: A numerical method. *arXiv preprint arXiv:2110.01343* (2021).
- [Lun18] A. LUNARDI. *Interpolation theory*, vol. 16 of *Appunti. Scuola Normale Superiore di Pisa (Nuova Serie) [Lecture Notes. Scuola Normale Superiore di Pisa (New Series)]*. Edizioni della Normale, Pisa, 2018. Third edition [of MR2523200]. <http://dx.doi.org/10.1007/978-88-7642-638-4>.
- [LZ22] C. LING and G. ZHAO. Nonlocal elliptic equation in Hölder space and the martingale problem. *J. Differential Equations* **314**, (2022), 653–699. <http://dx.doi.org/10.1016/j.jde.2022.01.025>.
- [MX18] R. MIKULEVIČIUS and F. XU. On the rate of convergence of strong Euler approximation for SDEs driven by Levy processes. *Stochastics* **90**, no. 4, (2018), 569–604. <http://dx.doi.org/10.1080/17442508.2017.1381095>.
- [MZ11] R. MIKULEVIČIUS and C. ZHANG. On the rate of convergence of weak euler approximation for nondegenerate SDEs driven by Lévy processes. *Stochastic Processes and their Applications* **121**, no. 8, (2011), 1720–1748. <http://dx.doi.org/10.1016/j.spa.2011.04.004>.
- [NO23] K. NASROALLAH and Y. OUKNINE. Euler approximation and stability of the solution to stochastic differential equations with jumps under pathwise uniqueness. *Stochastics* **95**, no. 2, (2023), 266–302. <http://dx.doi.org/10.1080/17442508.2022.2071107>.
- [Pap08] A. PAPAPANTOLEON. An introduction to Levy processes with applications in finance. *arXiv preprint arXiv:0804.0482* (2008).
- [Pri12] E. PRIOLA. Pathwise uniqueness for singular SDEs driven by stable processes. *Osaka Journal of Mathematics* **49**, no. 2, (2012), 421 – 447.
- [Pri15] E. PRIOLA. Stochastic flow for SDEs with jumps and irregular drift term. In *Stochastic analysis*, vol. 105 of *Banach Center Publ.*, 193–210. Polish Acad. Sci. Inst. Math., Warsaw, 2015. <http://dx.doi.org/10.4064/bc105-0-12>.
- [Pri18] E. PRIOLA. Davie’s type uniqueness for a class of SDEs with jumps. *Ann. Inst. Henri Poincaré Probab. Stat.* **54**, no. 2, (2018), 694–725. <http://dx.doi.org/10.1214/16-AIHP818>.

- [PS21] P. PRZYBYŁOWICZ and M. SZÖLGYENYI. Existence, uniqueness, and approximation of solutions of jump-diffusion sdes with discontinuous drift. *Applied Mathematics and Computation* **403**, (2021), 126191. <http://dx.doi.org/https://doi.org/10.1016/j.amc.2021.126191>.
- [PT17] O. M. PAMEN and D. TAGUCHI. Strong rate of convergence for the Euler–Maruyama approximation of SDEs with Hölder continuous drift coefficient. *Stochastic Processes and their Applications* **127**, no. 8, (2017), 2542 – 2559. <http://dx.doi.org/https://doi.org/10.1016/j.spa.2016.11.008>.
- [Sat13] K.-I. SATO. *Lévy processes and infinitely divisible distributions*, vol. 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2013. Translated from the 1990 Japanese original, Revised edition of the 1999 English translation.
- [SSW12] R. L. SCHILLING, P. SZTONYK, and J. WANG. Coupling property and gradient estimates of Lévy processes via the symbol. *Bernoulli* **18**, no. 4, (2012), 1128–1149. <http://dx.doi.org/10.3150/11-BEJ375>.
- [Tan92] T. TANIGUCHI. Successive approximations to solutions of stochastic differential equations. *J. Differential Equations* **96**, no. 1, (1992), 152–169. [http://dx.doi.org/10.1016/0022-0396\(92\)90148-G](http://dx.doi.org/10.1016/0022-0396(92)90148-G).
- [Yam81] T. YAMADA. On the successive approximation of solutions of stochastic differential equations. *J. Math. Kyoto Univ.* **21**, no. 3, (1981), 501–515. <http://dx.doi.org/10.1215/kjm/1250521975>.