The Annals of Applied Probability
2024, Vol. 34, No. 5, 4595–4634
https://doi.org/10.1214/24-AAP2075
© Institute of Mathematical Statistics, 2024

APPROXIMATION OF STOCHASTIC INTEGRALS WITH JUMPS VIA WEIGHTED BMO APPROACH

BY NGUYEN TRAN THUAN[a]

Department of Mathematics, Saarland University, [a]*nguyen@math.uni-saarland.de*

This article investigates discrete-time approximations of stochastic integrals driven by semimartingales with jumps via weighted bounded mean oscillation (BMO) approach. This approach enables L_p-estimates, $p \in (2, \infty)$, for the approximation error depending on the weight, and it allows a change of the underlying measure which leaves the error estimates unchanged. To take advantage of this approach, we propose a new approximation scheme obtained from an adjustment for the Riemann approximation based on tracking jumps of the underlying semimartingale. We discuss a way to optimize the approximation and also illustrate the sharpness of the obtained convergence rates. When the small jump activity of the semimartingale behaves like an α-stable process with $\alpha \in (1, 2)$, our scheme achieves under a regular regime the same convergence rate for the error as in Rosenbaum and Tankov [*Ann. Appl. Probab.* **24** (2014) 1002–1048]. Moreover, our approach extends to the case $\alpha \in (0, 1]$ and to the L_p-setting which are not treated there. As an application, we apply the methods in the special case where the semimartingale is an exponential Lévy process to mean-variance hedging of European type options.

1. Introduction.

1.1. *The problem and main results.* This article deals with discrete-time approximation problems for stochastic integrals and studies the error process $E = (E_t)_{t\in[0,T]}$ defined by

$$E_t := \int_0^t \vartheta_{u-}\,\mathrm{d}S_u - A_t, \tag{1.1}$$

where the time horizon $T \in (0, \infty)$ is fixed, ϑ is an admissible integrand, S is a semimartingale on a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ and $A = (A_t)_{t\in[0,T]}$ is an approximation scheme for the stochastic integral.

We will consider two approximation methods, where the second builds on the first one. For the first method, the *basic approximation method*, we assume that $A = A^{\mathrm{Rm}}$ is the Riemann approximation process of the above integral,

$$A_t^{\mathrm{Rm}}(\vartheta, \tau) := \sum_{i=1}^n \vartheta_{t_{i-1}-}(S_{t_i\wedge t} - S_{t_{i-1}\wedge t})$$

for the *deterministic* time-net $\tau_n = \{0 = t_0 < t_1 < \cdots < t_n = T\}$. We will study the corresponding error E^{Rm} in L_2, but *locally in time* in the sense that for any stopping time ρ with values in $[0, T]$ we measure the error which accumulates within $[\rho, T]$. The term *locally in time* also includes that at the random time ρ we restrict our problem to all sets $B \in \mathcal{F}_\rho$ of

Received December 2021; revised July 2023.
MSC2020 subject classifications. Primary 60H05, 91G20; secondary 60H07, 60G51.
Key words and phrases. Approximation of stochastic integral, convergence rate, Lévy process, semimartingale, weighted bounded mean oscillation.

positive measure, which leads to the notion of *bounded mean oscillation* (there are two abbreviations for it used in this article, bmo and BMO, which express two different spaces). Namely, we will work with *weighted* bmo-norms introduced in [16, 17] as we consider

$$\mathbb{E}^{\mathcal{F}_\rho}\big[|E_T^{\mathrm{Rm}} - E_\rho^{\mathrm{Rm}}|^2\big] \le c_{(1.2)}^2 \Phi_\rho^2 \quad \text{a.s., } \forall \rho. \tag{1.2}$$

Here, $\mathbb{E}^{\mathcal{F}_\rho}$ stands for the conditional expectation with respect to $\mathcal{F}_\rho$, and the *weight process* $\Phi = (\Phi_t)_{t\in[0,T]}$ will be specified later. Denote by $\|E^{\mathrm{Rm}}\|_{\mathrm{bmo}_2^\Phi(\mathbb{P})}$ the infimum of the $c_{(1.2)} > 0$ such that (1.2) is satisfied. We assert in Proposition 3.6 that, under certain conditions, one has

$$\|E^{\mathrm{Rm}}\|_{\mathrm{bmo}_2^\Phi(\mathbb{P})} \le c\sqrt{\|\tau_n\|_\theta},$$

where $\theta \in (0, 1]$ is related (but not only) to the growth property of the integrand ϑ by

$$\sup_{t\in[0,T)} (T-t)^{\frac{1-\theta}{2}} |\vartheta_t| < \infty \quad \text{a.s.,} \tag{1.3}$$

and $\|\tau_n\|_\theta$ denotes a nonlinear mesh size of τ_n related to θ. In Section 3.4 we discuss that τ_n can be chosen such that $\|\tau_n\|_\theta \le c/n$, implying the approximation rate

$$\|E^{\mathrm{Rm}}\|_{\mathrm{bmo}_2^\Phi(\mathbb{P})} \le cn^{-\frac{1}{2}}.$$

Roughly speaking, the faster the integrand grows as $t \uparrow T$, the more the time-net should be concentrated near T to compensate the growth.

If the semimartingale S has jumps, replacing E_ρ by $E_{\rho-}$ in (1.2) leads to different norms, the $\mathrm{BMO}_2^\Phi(\mathbb{P})$-norms. We will see in (1.9) and Proposition 2.5 that the $\mathrm{BMO}_2^\Phi(\mathbb{P})$-norm gives us a way to achieve good distributional tail estimates for the error E such as polynomial or exponential tail decay depending on the weight. Moreover, this approach allows us to switch the underlying measure $\mathbb{P}$ to an equivalent measure $\mathbb{Q}$, which is frequently encountered in mathematical finance, provided the change of measure satisfies a reverse Hölder inequality, so that the $\mathrm{BMO}_2^\Phi(\mathbb{Q})$-norm is equivalent to the $\mathrm{BMO}_2^\Phi(\mathbb{P})$-norm.

However, Example 3.7 below shows that if S has jumps, then the Riemann approximation error E^{Rm} does in general not converge to zero if measured in the $\mathrm{BMO}_2^\Phi(\mathbb{P})$-norm. The reason for this fact is that the $\mathrm{BMO}_2^\Phi(\mathbb{P})$-norm is relatively strong so that deterministic discretization times are not suitable to deal with possibly large jumps of S, which is in contrast to the case of no jump in [16]. To overcome this difficulty, we adapt and develop further the idea exploiting a small-large jump decomposition of S, which is used in the context of SDE discretization, see, for example, [10, 26–28], to our problem. This lets us design a new approximation scheme based on an adjustment of the Riemann sum which approximates the stochastic integral. This will be our second method, the *jump-adapted approximation method*, see Definition 3.9.

Generally speaking, the jump-adapted approximation $A^{\mathrm{adap}}(\vartheta, \tau|\varepsilon, \kappa)$ is of the form

$$A_t^{\mathrm{adap}}(\vartheta, \tau|\varepsilon, \kappa) = A_t^{\mathrm{Rm}}(\vartheta, \tau) + \mathrm{Correction}_t(\varepsilon, \kappa),$$

where the parameters $\varepsilon > 0$ and $\kappa \ge 0$ in the correction term relate to the threshold for which we decide which jumps of S are (relatively) large or small, and this threshold might continuously shrink when the time t approaches T, see Definition 3.8. The time-net used in this approximation method is a combination of the given deterministic time-net τ in the Riemann sum and random times of carefully chosen large jumps of S. A consequence of Proposition 3.15 shows that the expected value of the cardinality of this combined time-net is, up to a multiplicative constant, comparable to the cardinality of τ.

This new approximation scheme can be interpreted in the context of mathematical finance as follows: Before trading, we arrange to use the Riemann approximation $A^{\mathrm{Rm}}(\vartheta, \tau)$ associated with a trading strategy ϑ along with a preselected deterministic trading dates represented by τ. During trading with that initial plan, as soon as the large jumps of S occur, we trade additionally with the amount given in the correction term. So our point is that the decision for additional (random) trading times is based on the jump sizes of the price process, which can be observed, not on tracking jumps of the trading strategy that needs to be computed.

Denote by E^{adap} the error caused from the approximation with the jump-adapted scheme. To formulate the result, we assume that S is given as the (strong) solution of

$$\mathrm{d}S_t = \sigma(S_{t-})\,\mathrm{d}Z_t,$$

with σ specified later, where Z is a square integrable semimartingale defined in Section 2.3. Then Theorem 3.16 implies that, for suitably chosen time-nets and corrections, and for a suitable weight $\overline{\Phi}$, it holds that

$$\|E^{\mathrm{adap}}\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \le c n^{-\frac{1}{2}} \tag{1.4}$$

under the condition that the random measure π_Z of the predictable semimartingale characteristics of Z satisfies that $\pi_Z(\mathrm{d}t, \mathrm{d}z) = \nu_t(\mathrm{d}z)\,\mathrm{d}t$ and that

$$\sup_{r\in(0,1)} \left\| (\omega, t) \mapsto \int_{r<|z|\le 1} z \nu_t(\omega, \mathrm{d}z) \right\|_{L_\infty(\Omega\times[0,T],\mathbb{P}\otimes\lambda)} < \infty, \tag{1.5}$$

where λ is the Lebesgue measure, and one has

$$\|E^{\mathrm{adap}}\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \le c \begin{cases} n^{-\frac{1}{2\alpha}} & \text{if } \alpha \in (1,2], \\ n^{-\frac{1}{2}}(1+\log n) & \text{if } \alpha = 1, \\ n^{-\frac{1}{2}} & \text{if } \alpha \in (0,1) \end{cases} \tag{1.6}$$

provided that

$$\sup_{r\in(0,1)} \left\| (\omega, t) \mapsto r^\alpha \int_{r<|z|\le 1} \nu_t(\omega, \mathrm{d}z) \right\|_{L_\infty(\Omega\times[0,T],\mathbb{P}\otimes\lambda)} < \infty. \tag{1.7}$$

Condition (1.5) aims to indicate a local symmetry of ν around the origin rather than the small jump intensity of Z which is described by (1.7).

Since the integrator Z and structure conditions imposed on the approximated stochastic integral to achieve (1.4) and (1.6) are quite general, those obtained convergence rates are in general not the best possible. We will show in Section 3.3 that one can drastically improve those convergence rates in the particular case when S is the Doléans–Dade exponential of a pure jump process Z where Z has independent increments. Namely, Theorem 3.16 asserts that, for the error E^{adap} as above and under certain structure conditions for the approximated stochastic integral, one has

$$\|E^{\mathrm{adap}}\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \le c \begin{cases} n^{-\frac{1}{\alpha}(1-\frac{1}{2}(1-\theta)(\alpha-1))} & \text{if (1.7) holds for } \alpha \in (1,2], \\ n^{-1}(1+\log n) & \text{if (1.7) holds for } \alpha = 1, \\ n^{-1} & \text{if (1.7) holds for } \alpha = 1 \text{ and (1.5) holds}, \\ n^{-1} & \text{if (1.7) holds for } \alpha \in (0,1). \end{cases} \tag{1.8}$$

In the case $\alpha \in (1,2]$ in (1.8), the parameter $\theta \in (0,1]$ relates to the growth of ϑ mentioned in (1.3), and the corresponding exponent satisfies $\frac{3}{2\alpha} - \frac{1}{2} < \frac{1}{\alpha}(1 - \frac{1}{2}(1-\theta)(\alpha-1)) \le \frac{1}{\alpha}$. Since $\frac{3}{2\alpha} - \frac{1}{2} \ge \frac{1}{2\alpha}$ for $\alpha \in (1,2]$, the rate obtained in (1.8) is better than that in (1.6).

We discuss a lower bound for the approximation errors in Proposition 3.18 and provide a situation in Example 3.19 illustrating that the obtained convergence rates in (1.6) and in (1.8) are sharp. Specifically, the obtained rates are optimal when $\alpha \in (0,2]\backslash\{1\}$, and are optimal up to a log-factor when $\alpha = 1$.

Furthermore, Theorem 3.16 also reveals that, if the weight $\overline{\Phi}$ is sufficiently regular, then the estimates (1.4), (1.6) and (1.8) hold true for the L_p-norm, $p \in (2, \infty)$, in place of the $\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})$-norm. In addition, the measure $\mathbb{P}$ can be substituted by a suitable equivalent probability measure $\mathbb{Q}$ while keeping those estimates unchanged.

The parameter n in (1.4), (1.6) and (1.8) refers to certain moments of the cardinality of the combined time-net used in the approximation. This cardinality represents in the context of mathematical finance the number of transactions performed in trading, see Remark 3.17.

As an application, we choose S to be an exponential Lévy process and measure the discretization error for stochastic integrals where the integrands are mean-variance hedging (MVH) strategies of European payoffs. To do this, we provide in Proposition 4.2 an explicit representation of the MVH strategy for a European payoff for which we do not require any regularity for payoff functions nor specific structures from the underlying Lévy process. This result is, to the best of our knowledge, new in this generality and it might have an independent interest.

Let us end this subsection by listing some examples taken from Corollary 4.7 showing convergence rates for E^{adap} under the $\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})$-norm in the exponential Lévy setting. Namely, we let $S = \mathrm{e}^X$ where X is a Lévy process without the Brownian part whose small jump intensity behaves like an α-stable process with $\alpha \in (0, 2)$, and let ϑ be the MVH strategy of a payoff $g(S_T)$. Then, for the European call/put option (or any Lipschitz g), the convergence rate is of order: n^{-1} if $\alpha \in (0, 1)$, $n^{-1}(1+\log n)$ if $\alpha = 1$, and $n^{-\frac{1}{\alpha}}$ if $\alpha \in (1, 2)$. For the binary option (or any bounded g), the order of convergence rate is: n^{-1} if $\alpha \in (0, 1)$, $n^{-1}(1 + \log n)$ if $\alpha = 1$, and $n^{-\frac{1}{\alpha}[1-\frac{1}{\alpha}(\alpha-1)^2]+\delta}$ (for any $0 < \delta < \frac{1}{2}(1 - \frac{1}{\alpha})(\frac{2}{\alpha} - 1)$) if $\alpha \in (1, 2)$. Moreover, if $\mathbb{E}\mathrm{e}^{pX_T} < \infty$ for some $p \in (2, \infty)$, then measuring E^{adap} in L_p yields the same rates casewise as above. Last, our results are valid for some powered call/put options obtained from an interpolation, in a sense, between the binary and the call/put option.

1.2. *Literature overview.* Besides its own mathematical interest and its application to numerical methods, the approximation of a stochastic integral has a direct motivation in mathematical finance. Let us briefly discuss this for the Black–Scholes model. Assume that the (discounted) price of a risky asset is modelled by a stochastic process S which solves the SDE $\mathrm{d}S_t = \sigma(S_t)\,\mathrm{d}W_t$, where W is a standard Brownian motion and the function σ satisfies a suitable condition. For a European type payoff $g(S_T)$ satisfying an integrability condition, it is known that $g(S_T) = \mathbb{E}g(S_T) + \int_0^T \partial_y G(t, S_t)\,\mathrm{d}S_t$, where $G(t, y) := \mathbb{E}(g(S_T)|S_t = y)$ is the option price function and $(\partial_y G(t, S_t))_{t\in[0,T)}$ is the so-called delta-hedging strategy. The stochastic integral in the representation of $g(S_T)$ above can be interpreted as the theoretical hedging portfolio which is rebalanced continuously. However, it is not feasible in practice because one can only readjust the portfolio finitely many times. This leads to a replacement of the stochastic integral by a discretized version which causes the discretization error.

The error represented by the difference between a stochastic integral and its discretization has been extensively analyzed in various contexts. It is usually studied in L_2 for which one can exploit the orthogonality to reduce the probabilistic setting to a "more deterministic" setting where the corresponding quadratic variation is employed instead of the original error. In the Wiener space, we refer, for example, to [13, 15, 22], where the error along with its convergence rates was examined. The weak convergence of the error was treated in [18, 22]. When the driving process is a continuous semimartingale, the convergence in the L_2-sense was studied in [12], and in the almost sure sense it was considered in [21].

In this article, we allow the semimartingale to jump since many important processes used in financial modelling are not continuous (see, e.g., [6, 32]), and the presence of jumps has a significant effect on the hedging errors. Moreover, models with jumps typically correspond to incomplete markets. This means that beside the error resulting from the impossibility of continuously rebalancing a portfolio, there is another hedging error due to the incompleteness of the market. The latter problem was studied in many works (see an overview in [33] and the references therein). The present article mainly focuses on the first type of hedging error. The discretization error was studied within Lévy models in the weak convergence sense in [35], in the L_2-sense in [4, 14], and for a general jump model under the L_2-setting in [30].

In general, the classical L_2-approach for the error yields a second-order polynomial decay for its distributional tail by Markov's inequality. If higher-order decays are needed, then the L_p-approach ($2 < p < \infty$) is considered as a natural choice, and this direction has been investigated for diffusions on the Wiener space in [19]. A remarkably different route given in [16] is that one can study the error in weighted BMO spaces. The main benefit of the weighted BMO-approach is a John–Nirenberg type inequality ([16], Corollary 1(ii)): *If the error process E belongs to $\mathrm{BMO}_p^\Phi(\mathbb{P})$ for some $p \in (0, \infty)$, where Φ is some weight function specified in Definition 2.1, then there are constants $c, d > 0$ such that for any stopping time $\rho \colon \Omega \to [0, T]$ and any $\alpha, \beta > 0$,*

$$\mathbb{P}\Big(\sup_{u\in[\rho,T]} |E_u - E_{\rho-}| > c\alpha\beta \Big| \mathcal{F}_\rho\Big) \le \mathrm{e}^{1-\alpha} + d\mathbb{P}\Big(\sup_{u\in[\rho,T]} \Phi_u > \beta \Big| \mathcal{F}_\rho\Big). \tag{1.9}$$

Obviously, if Φ has a good distributional tail estimate, for example, if it has a polynomial or exponential tail decay, then by adjusting α and β one can derive a tail estimate for E accordingly. Especially, one can then derive L_p-estimates, $p \in (2, \infty)$, for the error. Some other applications of weighted BMO in the context of BSDEs have been considered in [20].

1.3. *Comparison to other works*. Regarding models with jumps, let us first mention the works done by Brodén and Tankov [4] and by Geiss, Geiss and Laukkarinen [14] which treat the Riemann approximation of stochastic integrals driven by the stochastic exponential of a Lévy process using deterministic discretization times. Although the approaches in [4] and [14] are different, both arrive at a result saying that, if the approximated stochastic integral is sufficiently regular, then the *asymptotically optimal* convergence rate of the error measured in L_2 is of order $n^{-\frac{1}{2}}$ when $n \to \infty$ (see [4], Corollary 3.1, and [14], Theorem 5), where n is the cardinality of the used time-net. For this direction, we also achieve in Theorem 3.16(1) the convergence rate of order $n^{-\frac{1}{2}}$ for the error under the bmo_2^Φ-norm, which then implies the same rate under the L_2-norm in our setting.

Later, Rosenbaum and Tankov [30] show that the convergence rate can be faster than $n^{-\frac{1}{2}}$ by using Riemann approximation associated with random discretization times. It is asserted in [30], Remark 5, that, when the semimartingale integrand does not possess a continuous local martingale part and if its small jump activity behaves like an α-stable process with $\alpha \in (1, 2)$, then the convergence rate measured in L_2 is of order $n^{-\frac{1}{\alpha}}$, which is also *asymptotically optimal* in their setting. In our framework, under a *regular regime* when $\theta = 1$, we derive from (1.8) the rate $n^{-\frac{1}{\alpha}}$ under the weighted BMO-norm (which is stronger than the L_2-norm in models that are originally only considered in the L_2-setting) when (1.7) holds for some $\alpha \in (1, 2]$. This rate is optimal in our setting and is consistent with that in [30], Remark 5, when n represents the expected number of transactions. Moreover, our results are valid for $\alpha \in (0, 1]$ which is not covered in [30].

We stress that our jump-adapted scheme is different from that in [30]. The authors in [30] use Riemann approximation schemes along with random times and the employed time-nets are the hitting times which are obtained by continuously tracking jumps of the integrand (which represents the trading strategy). Differently from that, time-nets in our method are a combination of preselected deterministic time-nets and random times obtained by continuously tracking jumps of the integrator (which represents the price process). In general, the computational cost of our method is less expensive than that in [30]. This can be argued in a situation when many options with different strategies are hedged at the same time with respect to a risky asset.

Other contributions of this work are, thanks to features of the (weighted) BMO-approach as aforementioned, to provide a situation that one can deduce L_p-estimates, $p \in (2, \infty)$, for the approximation error which are, to the best of our knowledge, still missing in the literature for models with jumps. Moreover, as a benefit to applications in mathematical finance, our results allow a change of the underlying measure which leaves the error estimates unchanged if the change of measure satisfies a reverse Hölder inequality, see Proposition 2.5.

1.4. *Structure of the article.* Some standard notions and notation are contained in Section 2. The main results are provided in Section 3 and their proofs are given in Section 5. In Section 4, we give some applications of those main results in exponential Lévy models. The regularity of weight processes used in this article is shown in Appendix B. Appendix C provides some gradient type estimates for a Lévy semigroup on Hölder spaces, which are used to verify the results in Section 4.

2. Preliminaries.

2.1. *Notation and conventions.*

General notation. Denote $\mathbb{N} := \{0, 1, 2, \dots\}$, $\mathbb{R}_+ := (0, \infty)$ and $\mathbb{R}_0 := \mathbb{R}\backslash\{0\}$. For $a, b \in \mathbb{R}$, we set $a \vee b := \max\{a, b\}$ and $a \wedge b := \min\{a, b\}$. For $A, B \geq 0$ and $c \geq 1$, the notation $A \sim_c B$ stands for $A/c \leq B \leq cA$. The notation log indicates the logarithm to the base 2 and $\log^+ x := \log(x \vee 1)$. Subindexing a symbol by a label means the place where that symbol appears (e.g., $c_{(2.1)}$ refers to the relation (2.1)).

The Lebesgue measure on the Borel σ-algebra $\mathcal{B}(\mathbb{R})$ is denoted by λ, and we also write $\mathrm{d}x$ instead of $\lambda(\mathrm{d}x)$ for simplicity. For $p \in [1, \infty]$ and $A \in \mathcal{B}(\mathbb{R})$, the notation $L_p(A)$ means the space of all p-order integrable Borel functions on A with respect to λ, where the essential supremum is taken when $p = \infty$.

Let ξ be a random variable defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The push-forward measure of $\mathbb{P}$ with respect to ξ is denoted by $\mathbb{P}_\xi$. If ξ is integrable (nonnegative), then the (generalized) conditional expectation of ξ given a sub-σ-algebra $\mathcal{G} \subseteq \mathcal{F}$ is denoted by $\mathbb{E}^{\mathcal{G}}[\xi]$. We also agree on the notation $L_p(\mathbb{P}) := L_p(\Omega, \mathcal{F}, \mathbb{P})$.

Notation for stochastic processes. Let $T \in (0, \infty)$ be fixed and $(\Omega, \mathcal{F}, \mathbb{P})$ a complete probability space equipped with a right continuous filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$. Assume that $\mathcal{F}_0$ is generated by $\mathbb{P}$-null sets only. Because of the conditions imposed on $\mathbb{F}$, we may assume that every martingale adapted to this filtration is *càdlàg* (right-continuous with left limits). For $\mathbb{I} = [0, T]$ or $\mathbb{I} = [0, T)$, we use the following notation:

- For two processes $X = (X_t)_{t\in\mathbb{I}}$, $Y = (Y_t)_{t\in\mathbb{I}}$, by writing $X = Y$ we mean that $X_t = Y_t$ for all $t \in \mathbb{I}$ a.s., and similarly when the relation "=" is replaced by some standard relations such as "≥", "≤", etc.
- For a càdlàg process $X = (X_t)_{t\in\mathbb{I}}$, we define the process $X_- = (X_{t-})_{t\in\mathbb{I}}$ by setting $X_{0-} := X_0$ and $X_{t-} := \lim_{0<s\uparrow t} X_s$ for $t \in \mathbb{I}\backslash\{0\}$. In addition, set $\Delta X := X - X_-$.
- $\mathrm{CL}(\mathbb{I})$ denotes the family of all càdlàg and $\mathbb{F}$-adapted processes $X = (X_t)_{t\in\mathbb{I}}$.
- $\mathrm{CL}_0(\mathbb{I})$ (resp. $\mathrm{CL}^+(\mathbb{I})$) consists of all $X \in \mathrm{CL}(\mathbb{I})$ with $X_0 = 0$ a.s. (resp. $X \geq 0$).
- Let $M = (M_t)_{t\in\mathbb{I}}$ and $N = (N_t)_{t\in\mathbb{I}}$ be $L_2(\mathbb{P})$-martingales adapted to $\mathbb{F}$. The *predictable quadratic covariation* of M and N is denoted by $\langle M, N\rangle$. If $M = N$, then we simply write $\langle M\rangle$ instead of $\langle M, M\rangle$.
- For $p \in [1, \infty]$ and $X \in \mathrm{CL}([0, T])$, we denote $\|X\|_{S_p(\mathbb{P})} := \|\sup_{t\in[0,T]} |X_t|\|_{L_p(\mathbb{P})}$.

2.2. *Weighted bounded mean oscillation and regular weight.* We recall the notions of weighted bounded mean oscillation and the space $\mathcal{SM}_p(\mathbb{P})$ of regular weight processes (the abbreviation $\mathcal{SM}$ indicates the property resembling a supermartingale). Let $\mathcal{S}([0,T])$ be the family of all stopping times $\rho\colon \Omega \to [0,T]$, and set $\inf\varnothing := \infty$.

DEFINITION 2.1 ([16, 17]). *For $p \in (0,\infty)$, $Y \in \mathrm{CL}_0([0,T])$ and $\Phi \in \mathrm{CL}^+([0,T])$, let*

$$\|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{P})} := \inf\{c \ge 0 : \mathbb{E}^{\mathcal{F}_\rho}[|Y_T - Y_{\rho-}|^p] \le c^p \Phi_\rho^p \text{ a.s.}, \forall \rho \in \mathcal{S}([0,T])\},$$

$$\|Y\|_{\mathrm{bmo}_p^\Phi(\mathbb{P})} := \inf\{c \ge 0 : \mathbb{E}^{\mathcal{F}_\rho}[|Y_T - Y_{\rho}|^p] \le c^p \Phi_\rho^p \text{ a.s.}, \forall \rho \in \mathcal{S}([0,T])\},$$

$$\|\Phi\|_{\mathcal{SM}_p(\mathbb{P})} := \inf\Big\{c \ge 0 : \mathbb{E}^{\mathcal{F}_\rho}\Big[\sup_{\rho\le t\le T} \Phi_t^p\Big] \le c^p \Phi_\rho^p \text{ a.s.}, \forall \rho \in \mathcal{S}([0,T])\Big\}.$$

For $\Gamma \in \{\mathrm{BMO}_p^\Phi(\mathbb{P}), \mathrm{bmo}_p^\Phi(\mathbb{P})\}$, if $\|Y\|_\Gamma < \infty$ (resp. $\|\Phi\|_{\mathcal{SM}_p(\mathbb{P})} < \infty$), then we write $Y \in \Gamma$ (resp. $\Phi \in \mathcal{SM}_p(\mathbb{P})$). In the nonweighted case, that is, $\Phi \equiv 1$, we drop Φ and simply use the notation $\mathrm{BMO}_p(\mathbb{P})$ or $\mathrm{bmo}_p(\mathbb{P})$.

REMARK 2.2. *According to [17], Propositions A.4 and A.1, the definitions of $\|\cdot\|_{\mathrm{bmo}_p^\Phi(\mathbb{P})}$ and $\|\cdot\|_{\mathcal{SM}_p(\mathbb{P})}$ can be simplified by using deterministic times $a \in [0,T]$ instead of stopping times ρ, that is,*

$$\|Y\|_{\mathrm{bmo}_p^\Phi(\mathbb{P})} = \inf\{c \ge 0 : \mathbb{E}^{\mathcal{F}_a}[|Y_T - Y_a|^p] \le c^p \Phi_a^p \text{ a.s.}, \forall a \in [0,T]\},$$

$$\|\Phi\|_{\mathcal{SM}_p(\mathbb{P})} = \inf\Big\{c \ge 0 : \mathbb{E}^{\mathcal{F}_a}\Big[\sup_{a\le t\le T} \Phi_t^p\Big] \le c^p \Phi_a^p \text{ a.s.}, \forall a \in [0,T]\Big\}.$$

The theory of classical nonweighted BMO/bmo-martingales can be found in [9], Chapter VII, or [29], Chapter IV, and they were used later in different contexts (see, e.g., [5, 8]). The notion of weighted BMO space above was introduced and discussed in [16] where it was developed for general càdlàg processes which are not necessarily martingales.

It is clear that if $Y \in \mathrm{CL}_0([0,T])$ is continuous, then $\|Y\|_{\mathrm{bmo}_p^\Phi(\mathbb{P})} = \|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{P})}$. If Y has jumps, then the relation between weighted BMO and weighted bmo is as follows.

LEMMA 2.3 ([17], Propositions A.5 and A.3). *If $\Phi \in \mathcal{SM}_p(\mathbb{P})$ for some $p \in (0,\infty)$, then there is a constant $c = c(p, \|\Phi\|_{\mathcal{SM}_p(\mathbb{P})}) > 0$ such that for all $Y \in \mathrm{CL}_0([0,T])$,*

$$\|Y\|_{\mathrm{BMO}_p^\Phi(\mathbb{P})} \sim_c \|Y\|_{\mathrm{bmo}_p^\Phi(\mathbb{P})} + |\Delta Y|_\Phi,$$

where $|\Delta Y|_\Phi := \inf\{c \ge 0 : |\Delta Y_t| \le c\Phi_t$ for all $t \in [0,T]$ a.s.$\}$.

DEFINITION 2.4 ([16]). *Let $\mathbb{Q}$ be an equivalent probability measure to $\mathbb{P}$ so that $U := \mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P} > 0$. Then $\mathbb{Q} \in \mathcal{RH}_s(\mathbb{P})$ for some $s \in (1,\infty)$ if $U \in L_s(\mathbb{P})$ and if there is a constant $c_{(2.1)} > 0$ such that U satisfies the following reverse Hölder inequality:*

$$\mathbb{E}^{\mathcal{F}_\rho}[U^s] \le c_{(2.1)}^s \big(\mathbb{E}^{\mathcal{F}_\rho}[U]\big)^s \quad \text{a.s.}, \forall \rho \in \mathcal{S}([0,T]), \tag{2.1}$$

where the conditional expectation $\mathbb{E}^{\mathcal{F}_\rho}$ is computed under $\mathbb{P}$.

We recall some features of weighted BMO which play a key role in application.

PROPOSITION 2.5 ([16, 17]). *Let $p, q \in (0,\infty)$ and $\Phi \in \mathrm{CL}^+([0,T])$.*

(1) *There is a constant $c_1 = c_1(p,q) > 0$ such that $\|\cdot\|_{S_p(\mathbb{P})} \le c_1 \|\Phi\|_{S_p(\mathbb{P})} \|\cdot\|_{\mathrm{BMO}_q^\Phi(\mathbb{P})}$.*

(2) *If* $\Phi \in \mathcal{SM}_p(\mathbb{P})$, *then for any* $r \in (0, p]$ *a constant* $c_2 = c_2(r, p, \|\Phi\|_{\mathcal{SM}_p(\mathbb{P})}) > 0$ *exists such that* $\|\cdot\|_{\mathrm{BMO}_p^\Phi(\mathbb{P})} \sim_{c_2} \|\cdot\|_{\mathrm{BMO}_r^\Phi(\mathbb{P})}$.

(3) *If* $\mathbb{Q} \in \mathcal{RH}_s(\mathbb{P})$ *for some* $s \in (1, \infty)$ *and* $\Phi \in \mathcal{SM}_p(\mathbb{Q})$, *then there is a constant* $c_3 = c(s, p, \mathbb{Q}, \Phi) > 0$ *such that* $\|\cdot\|_{\mathrm{BMO}_p^\Phi(\mathbb{Q})} \le c_3 \|\cdot\|_{\mathrm{BMO}_p^\Phi(\mathbb{P})}$.

PROOF. Items (1) and (2) are due to [17], Proposition A.6. For Item (3), we apply [16], combine Corollary 1(i) with Theorem 3, to the weight $\Phi + \varepsilon > 0$ and then let $\varepsilon \downarrow 0$. □

2.3. *The class of approximated stochastic integrals.* Throughout this article, the assumptions for the stochastic integral in (1.1) are the following.

[S] The process $S \in \mathrm{CL}([0, T])$ is a strong solution of the SDE[1]

$$\mathrm{d}S_t = \sigma(S_{t-})\,\mathrm{d}Z_t, \quad S_0 \in \mathcal{R}_S, \tag{2.2}$$

where $\sigma\colon \mathcal{R}_S \to (0, \infty)$ is a Lipschitz function on an open set $\mathcal{R}_S \subseteq \mathbb{R}$ with $S_t(\omega)$, $S_{t-}(\omega) \in \mathcal{R}_S$ for all $(\omega, t) \in \Omega \times [0, T]$. We denote

$$|\sigma|_{\mathrm{Lip}} := \sup_{x, y \in \mathcal{R}_S, x \neq y} \left| \frac{\sigma(y) - \sigma(x)}{y - x} \right| < \infty.$$

[Z] The process $Z \in \mathrm{CL}([0, T])$ is an $L_2(\mathbb{P})$-semimartingale on $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ with the representation

$$Z_t = Z_0 + Z_t^{\mathrm{c}} + \int_0^t \int_{\mathbb{R}_0} z(N_Z - \pi_Z)(\mathrm{d}u, \mathrm{d}z) + \int_0^t b_u^Z\,\mathrm{d}u, \quad t \in [0, T], \tag{2.3}$$

where $Z_0 \in \mathbb{R}$, b^Z is a progressively measurable process, Z^{c} is a pathwise continuous $L_2(\mathbb{P})$-martingale with $Z_0^{\mathrm{c}} = 0$, N_Z is the jump random measure[2] of Z and π_Z is the predictable compensator[3] of N_Z. Assumptions on Z are the following:

(Z1) For all $\omega \in \Omega$,

$$\pi_Z(\omega, \mathrm{d}t, \mathrm{d}z) = \nu_t(\omega, \mathrm{d}z)\,\mathrm{d}t, \tag{2.4}$$

where the transition kernel[4] $\nu_t(\omega, \cdot)$ is a Lévy measure, that is, a Borel measure on $\mathcal{B}(\mathbb{R})$ satisfying $\nu_t(\omega, \{0\}) := 0$ and $\int_{\mathbb{R}}(z^2 \wedge 1)\nu_t(\omega, \mathrm{d}z) < \infty$.

(Z2) There is a progressively measurable process a^Z such that $\mathrm{d}\langle Z^{\mathrm{c}}\rangle_t = |a_t^Z|^2\,\mathrm{d}t$ and

$$a_{(2.5)}^Z := \|a^Z\|_{L_\infty(\Omega\times[0,T], \mathbb{P}\otimes\lambda)} < \infty. \tag{2.5}$$

(Z3) The processes b^Z and j^Z, where $j_t^Z := (\int_{\mathbb{R}} z^2 \nu_t(\mathrm{d}z))^{1/2}$, satisfy that

$$b_{(2.6)}^Z := \big\| \|b^Z\|_{L_2([0,T],\lambda)} \big\|_{L_\infty(\mathbb{P})} < \infty, \qquad j_{(2.6)}^Z := \|j^Z\|_{L_\infty(\Omega\times[0,T], \mathbb{P}\otimes\lambda)} < \infty. \tag{2.6}$$

[I] The process ϑ belongs to the family $\mathcal{A}(S)$ of *admissible integrands*, where

$$\mathcal{A}(S) := \left\{ \vartheta \in \mathrm{CL}([0, T)) : \mathbb{E} \int_0^T \vartheta_{t-}^2 \sigma(S_{t-})^2\,\mathrm{d}t < \infty \text{ and } \Delta\vartheta_t = 0 \text{ a.s.}, \forall t \in [0, T) \right\}.$$

[1] See, for example, [29], Chapter V, Section 3, for the existence and uniqueness of S.

[2] $N_Z((s, t] \times B) := \#\{u \in (s, t] : \Delta Z_u \in B\}$ and $N_Z(\{0\} \times B) := 0$ for $0 \le s < t \le T$, $B \in \mathcal{B}(\mathbb{R}_0)$.

[3] π_Z is such that: (i) for any $\omega \in \Omega$, $\pi_Z(\omega, \cdot)$ is a measure on $\mathcal{B}([0, T] \times \mathbb{R})$ with $\pi_Z(\omega, \{0\} \times \mathbb{R}) = 0$; (ii) for any $\mathcal{P} \otimes \mathcal{B}(\mathbb{R})$-measurable and nonnegative f, the process $\int_0^\cdot \int_{\mathbb{R}} f(u, z)\pi_Z(\mathrm{d}u, \mathrm{d}z)$ is $\mathcal{P}$-measurable satisfying $\mathbb{E}\int_0^T \int_{\mathbb{R}} f(u, z) N_Z(\mathrm{d}u, \mathrm{d}z) = \mathbb{E}\int_0^T \int_{\mathbb{R}} f(u, z)\pi_Z(\mathrm{d}u, \mathrm{d}z)$, where $\mathcal{P}$ is the predictable σ-algebra on $\Omega \times [0, T]$ (see [25], Chapter II, Section 1, for more details).

[4] In the sense of [25], Chapter II, Theorem 1.8.

REMARK 2.6.

(1) *By a standard stopping argument and Gronwall's lemma*, (2.2) *implies that S is an $L_2(\mathbb{P})$-semimartingale and*

$$\mathbb{E}\int_0^T \sigma(S_u)^2\,\mathrm{d}u = \mathbb{E}\int_0^T \sigma(S_{u-})^2\,\mathrm{d}u < \infty.$$

(2) *For each $t \in [0,T]$, it follows from* (2.4) *that $N_Z(\{t\}\times\mathbb{R}_0) = 0$ a.s., which verifies $\Delta Z_t = 0$ a.s., and hence, $\Delta S_t = 0$ a.s. In other words, Z and S have no fixed-time discontinuity. Since admissible integrands ϑ in applications are often functionals of the integrator S, it is technically convenient to assume that $\Delta\vartheta_t = 0$ a.s.*

3. Approximation via weighted bounded mean oscillation approach. To examine the discrete-time approximation problem in weighted bmo or weighted BMO, further structure of the integrand is required. We begin with the following assumption which is an adaptation of [17], Assumption 5.1.

ASSUMPTION 3.1. *Assume for a $\vartheta \in \mathcal{A}(S)$ that there exists a random measure*

$$\Upsilon : \Omega \times \mathcal{B}((0,T)) \to [0,\infty]$$

such that $\Upsilon(\omega,(0,t]) < \infty$ for all $(\omega,t) \in \Omega\times(0,T)$ and there is a constant $c_{(3.1)} > 0$ such that for any $0 \le a < b < T$, a.s.,

$$\mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,b]} |\vartheta_t - \vartheta_a|^2\sigma(S_t)^2\,\mathrm{d}t\right] \le c_{(3.1)}^2\mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,b]}(b-t)\Upsilon(\cdot,\mathrm{d}t)\right]. \tag{3.1}$$

The left-hand side of (3.1) appears as the one-step conditional L_2-approximation error. This error is assumed to be controlled from above by a conditional integral with respect to an appropriate random measure Υ, where Υ might have some singularity at the terminal time T. This structure condition allows us to derive the multi-step approximation from the one-step approximation. Apparently, the measure Υ looks artificial, however, it origins from the diffusion setting on the Wiener space. We briefly explain this in Example 3.2 below.

EXAMPLE 3.2 (Diffusion setting). *We recall the setting from* [13] (*see also* [17]): *Let $\hat{\sigma} : \mathbb{R} \to \mathbb{R}$ be bounded, $\inf_{x\in\mathbb{R}}\hat{\sigma}(x) > 0$ and infinitely differentiable with bounded derivatives. Let $\sigma(x) := x\hat{\sigma}(\ln x)$, which is Lipschitz on $(0,\infty)$, and consider the SDE*

$$\mathrm{d}S_t = \sigma(S_t)\,\mathrm{d}W_t, \qquad S_0 = \mathrm{e}^{x_0} > 0,$$

where W is a standard Brownian motion. Then one has $S_t = \mathrm{e}^{X_t}$, where $\mathrm{d}X_t = \hat{\sigma}(X_t)\,\mathrm{d}W_t - \frac{1}{2}\hat{\sigma}(X_t)^2\,\mathrm{d}t$, $X_0 = x_0 \in \mathbb{R}$. For any Borel function $g : (0,\infty) \to \mathbb{R}$ with polynomial growth, Itô's formula asserts $g(S_T) = \mathbb{E}g(S_T) + \int_0^T \partial_y G(t,S_t)\,\mathrm{d}S_t$ with $G(t,y) := \mathbb{E}(g(S_T)|S_t = y)$. Then [15], *Corollary* 3.3, *verifies that Assumption* 3.1 *is satisfied for*

$$\vartheta_t := \partial_y G(t,S_t), \qquad \Upsilon(\omega,\mathrm{d}t) := \left|(\sigma^2\partial_y^2 G)(t,S_t(\omega))\right|^2\mathrm{d}t.$$

In this case, since Υ is related to the second derivative, it describes some kind of curvature of the stochastic integral.

We now provide in Example 3.3 another formula for Υ which is used later in the exponential Lévy setting in Section 4.

EXAMPLE 3.3. *Assume for $\vartheta \in \mathcal{A}(S)$ that $\vartheta\sigma(S)$ has the representation*

$$\vartheta_t\sigma(S_t) = M_t + V_t, \quad t \in [0, T),$$

where $M = (M_t)_{t\in[0,T)}$ is an $L_2(\mathbb{P})$-martingale, $V_t = \int_0^t v_u \,\mathrm{d}u$ for a progressively measurable v with $\int_0^t v_u^2(\omega)\,\mathrm{d}u < \infty$ for all $(\omega, t) \in \Omega \times [0, T)$. Then Assumption 3.1 is satisfied for

$$\Upsilon(\omega, \mathrm{d}t) := \mathrm{d}\langle M\rangle_t(\omega) + |\sigma|_{\mathrm{Lip}}^2 M_t^2(\omega)\,\mathrm{d}t + v_t^2(\omega)\,\mathrm{d}t + |\sigma|_{\mathrm{Lip}}^2\left(\int_0^t v_u^2(\omega)\,\mathrm{d}u\right)\mathrm{d}t.$$

The proof for this assertion is provided in the Supplementary Material [37], *subsection D.2.*

The key assumption which enables to derive the approximation results is as follows.

ASSUMPTION 3.4. *Let $\theta \in (0, 1]$. Assume that Assumption 3.1 is satisfied, and there are processes $\Theta, \Phi \in \mathrm{CL}^+([0, T])$, where Θ is a.s. nondecreasing with $\Theta\sigma(S) \le \Phi$, such that the following two conditions hold*:

(a) (*Growth condition*) *There is a constant $c_{(3.2)} > 0$ such that*

$$|\vartheta_a| \le c_{(3.2)}(T - a)^{\frac{\theta-1}{2}}\Theta_a \quad a.s., \forall a \in [0, T). \tag{3.2}$$

(b) (*Curvature condition*) *There is a constant $c_{(3.3)} > 0$ such that*

$$\mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,T)} (T - t)^{1-\theta}\Upsilon(\cdot, \mathrm{d}t)\right] \le c_{(3.3)}^2\Phi_a^2 \quad a.s., \forall a \in [0, T). \tag{3.3}$$

Here, the constants $c_{(3.2)}$ and $c_{(3.3)}$ may depend on θ.

The parameter θ in (3.2) describes the growth (pathwise and relatively to Θ) of ϑ and the integrand $(T - t)^{1-\theta}$ in (3.3) is employed to compensate the singularity of Υ when the time variable approaches T. Hence, the bigger θ is, the less singular at T of both ϑ and Υ get, which leads the approximated stochastic integral to be more regular. In particular, for the Black–Scholes model with the delta-hedging strategy ϑ, that is, $\sigma(x) = x$ in Example 3.2, the parameter θ can be interpreted as the *fractional smoothness* of the payoff in the sense of [13, 19] where $\theta = 1$ corresponds to the smoothness of order 1. Thus, in this article we will refer to the situation when Assumption 3.4 holds for $\theta = 1$ as *regular regime*. It is also clear that if Assumption 3.4 is satisfied for a $\theta \in (0, 1]$, then it also holds for any $\theta' \in (0, \theta)$ with the same Θ and Φ and with a suitable change for $c_{(3.2)}$ and $c_{(3.3)}$.

Various specifications of Assumption 3.4 in the Brownian setting or in the Lévy setting are provided in [17]. In Section 4 below, we consider Assumption 3.4 in the exponential Lévy setting which extends [17].

3.1. *The basic method*: *Riemann approximation.*

DEFINITION 3.5.

(1) *Let $\mathcal{T}_{\det}$ be the family of all deterministic time-nets $\tau = (t_i)_{i=0}^n$ on $[0, T]$ with $0 = t_0 < t_1 < \cdots < t_n = T$, $n \ge 1$. The mesh size of $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\det}$ is measured with respect to a parameter $\theta \in (0, 1]$ by*

$$\|\tau\|_\theta := \max_{i=1,\dots,n} \frac{t_i - t_{i-1}}{(T - t_{i-1})^{1-\theta}}.$$

(2) *For* $\vartheta \in \mathcal{A}(S)$, $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\det}$ *and* $t \in [0, T]$, *we define*

$$A_t^{\mathrm{Rm}}(\vartheta, \tau) := \sum_{i=1}^{n} \vartheta_{t_{i-1}-}(S_{t_i \wedge t} - S_{t_{i-1} \wedge t}), \qquad E_t^{\mathrm{Rm}}(\vartheta, \tau) := \int_0^t \vartheta_{u-}\, \mathrm{d}S_u - A_t^{\mathrm{Rm}}(\vartheta, \tau).$$

The following is based on [17], Theorem 5.3, and its proof is presented in Section 5.1.

PROPOSITION 3.6. *Let Assumption 3.4 hold for some* $\theta \in (0, 1]$. *Then there exists a constant* $c_{(3.4)} > 0$ *such that for any* $\tau \in \mathcal{T}_{\det}$,

$$\|E^{\mathrm{Rm}}(\vartheta, \tau)\|_{\mathrm{bmo}_2^{\Phi}(\mathbb{P})} \le c_{(3.4)} \sqrt{\|\tau\|_\theta}. \tag{3.4}$$

In particular, if S *is continuous then* $\|E^{\mathrm{Rm}}(\vartheta, \tau)\|_{\mathrm{BMO}_2^{\Phi}(\mathbb{P})} \le c_{(3.4)} \sqrt{\|\tau\|_\theta}$ *for any* $\tau \in \mathcal{T}_{\det}$.

3.2. *The jump-adapted approximation*: *General results*. In Proposition 3.6, the continuity of S is crucial to derive the conclusion for $E^{\mathrm{Rm}}(\vartheta, \tau)$ under the $\mathrm{BMO}_2^{\Phi}(\mathbb{P})$-norm. If S has jumps, then that result may fail as shown in the following example.

EXAMPLE 3.7. *In the notation of Section 2.3, we let* $Z = \tilde{J}$, *where* $\tilde{J}_t := J_t - rt$ *is a compensated Poisson process with intensity* $r > 0$. *Choose* $\sigma \equiv 1$ *so that* $S = Z$. *Let* $f : (0, T] \times \mathbb{N} \to \mathbb{R}$ *be a Borel function with* $\|f\|_\infty := \sup_{(t,k) \in (0,T] \times \mathbb{N}} |f(t, k)| < \infty$ *and* $\varepsilon := \inf_{t \in (0,T]} |f(t, 0)| > 0$. *Assume that*

$$\delta := \varepsilon - rT\|f\|_\infty > 0.$$

Let $\rho_1 := \inf\{t > 0 : \Delta J_t = 1\} \wedge T$ *and* $\rho_2 := \inf\{t > \rho_1 : \Delta J_t = 1\} \wedge T$. *Let* $\vartheta_0 \in \mathbb{R}$ *and define* $\vartheta_t = \vartheta_0 + \int_{(0,t\wedge\rho_2]} f(s, J_{s-})\, \mathrm{d}\tilde{J}_s$, $t \in (0, T]$. *It is not difficult to check that* $\vartheta \in \mathcal{A}(S)$ *is a martingale with* $\|\vartheta_T\|_{L_\infty(\mathbb{P})} < \infty$. *Then Assumption 3.1 is satisfied with the selection* $\Upsilon(\cdot, \mathrm{d}t) := \mathrm{d}\langle\vartheta\rangle_t$ *as showed in Example 3.3* (*with* $V \equiv 0$). *In addition, it is straightforward to check that Assumption 3.4 holds true for* $\Theta \equiv \Phi \equiv 1$ *and for* $\theta = 1$.

Take $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\det}$ *arbitrarily. On the set* $\{0 < \rho_1 < \rho_2 < t_1\}$ *we have*

$$\begin{aligned}
|\Delta E_{\rho_2}^{\mathrm{Rm}}(\vartheta, \tau)| &= \sum_{i=1}^{n} |\vartheta_{\rho_2-} - \vartheta_{t_{i-1}-}| \mathbb{1}_{(t_{i-1},t_i]}(\rho_2) |\Delta J_{\rho_2}| = |\vartheta_{\rho_2-} - \vartheta_0| \\
&= \left| f(\rho_1, J_{\rho_1-}) - r \int_{(0,\rho_2)} f(s, J_{s-})\, \mathrm{d}s \right| \ge |f(\rho_1, 0)| - rT\|f\|_\infty \\
&\ge \delta.
\end{aligned}$$

Since $\mathbb{P}(0 < \rho_1 < \rho_2 < t_1) > 0$, *it implies that* $\inf_{\tau \in \mathcal{T}_{\det}} \|\Delta E_{\rho_2}^{\mathrm{Rm}}(\vartheta, \tau)\|_{L_\infty(\mathbb{P})} \ge \delta$. *Due to Lemma 2.3, we obtain* $\inf_{\tau \in \mathcal{T}_{\det}} \|E^{\mathrm{Rm}}(\vartheta, \tau)\|_{\mathrm{BMO}_p(\mathbb{P})} > 0$ *for any* $p \in (0, \infty)$.

Therefore, in order to exploit benefits of the weighted BMO approach, we introduce a new approximation scheme based on an adjustment of the classical Riemann approximation. The time-net for this scheme is obtained by combining a given deterministic time-net, which is used in the Riemann sum of the stochastic integral, and a suitable sequence of random times which captures the (relative) large jumps of the driving process. With this scheme, we not only can utilize the features of weighted BMO, but can also control the cardinality of the combined time-nets.

Let us begin with the random times. Due to the assumptions imposed on S in Section 2.3, one has $\sigma(S_-) > 0$ and

$$\Delta S = \sigma(S_-)\Delta Z \tag{3.5}$$

from which we can see that jumps of S can be determined from knowing jumps of Z. However, if we would use S to model the stock price process, then it is more realistic to track the jumps of S rather than of Z. Therefore, we define the random times $\rho(\varepsilon,\kappa)=(\rho_i(\varepsilon,\kappa))_{i\geq 0}$ based on tracking the jumps of S as follows (recall that $\inf\varnothing:=\infty$).

DEFINITION 3.8. *For $\varepsilon>0$ and $\kappa\geq 0$, let $\rho_0(\varepsilon,\kappa):=0$ and*

$$\rho_i(\varepsilon,\kappa):=\inf\{T\geq t>\rho_{i-1}(\varepsilon,\kappa):|\Delta S_t|>\sigma(S_{t-})\varepsilon(T-t)^\kappa\}\wedge T,\quad i\geq 1, \tag{3.6}$$

$$\mathcal{N}_{(3.7)}(\varepsilon,\kappa):=\inf\{i\geq 1:\rho_i(\varepsilon,\kappa)=T\}. \tag{3.7}$$

The quantity $\varepsilon(T-t)^\kappa$ in (3.6) is the level at time t from which we decide which jumps of S are (relatively) large, and moreover, for $\kappa>0$, this level continuously shrinks when t approaches the terminal time T. Hence, κ describes the *jump size decay rate*. The idea for using the decay function $(T-t)^\kappa$ is to compensate the growth of integrands. By specializing $T=1$ and letting $t=0$, the parameter ε can be interpreted as the *initial jump size threshold*.

DEFINITION 3.9 (Jump-adapted approximation). *Let $\varepsilon>0$, $\kappa\in[0,\frac{1}{2})$ and $\tau\in\mathcal{T}_{\mathrm{det}}$.*

(1) *Let $\tau\sqcup\rho(\varepsilon,\kappa)$ be the (random) discretization times of $[0,T]$ by combining τ with $\rho(\varepsilon,\kappa)$ and re-ordering their time-knots.*

(2) *For $\tau=(t_i)_{i=0}^n$ and $t\in[0,T]$, we set $\vartheta_t^\tau:=\sum_{i=1}^n\vartheta_{t_{i-1}-}\mathbb{1}_{(t_{i-1},t_i]}(t)$ and define*

$$\begin{aligned} A_t^{\mathrm{adap}}(\vartheta,\tau|\varepsilon,\kappa)&:=A_t^{\mathrm{Rm}}(\vartheta,\tau)+\sum_{\rho_i(\varepsilon,\kappa)\in[0,t]\cap[0,T)}\big(\vartheta_{\rho_i(\varepsilon,\kappa)-}-\vartheta^\tau_{\rho_i(\varepsilon,\kappa)}\big)\Delta S_{\rho_i(\varepsilon,\kappa)},\\ E_t^{\mathrm{adap}}(\vartheta,\tau|\varepsilon,\kappa)&:=\int_0^t\vartheta_{u-}\,\mathrm{d}S_u-A_t^{\mathrm{adap}}(\vartheta,\tau|\varepsilon,\kappa),\end{aligned} \tag{3.8}$$

where $A^{\mathrm{Rm}}(\vartheta,\tau)$ is given in Definition 3.5.

As verified later in Section 5.2, each $\rho_i(\varepsilon,\kappa)$ is a stopping time. Moreover, the sum on the right-hand side of (3.8) is a finite sum a.s. as a consequence of Lemma 5.3 below. Hence, by adjusting this sum on a set of probability zero, we may assume that $A^{\mathrm{adap}}(\vartheta,\tau|\varepsilon,\kappa)\in\mathrm{CL}_0([0,T])$. Besides, we also restrict the sum over the stopping times taking values in $[0,T)$ instead of $[0,T]$ because of two technical reasons. First, ϑ does not necessarily have the left-limit at T, and second, since $\Delta S_T=0$ a.s. as mentioned in Remark 2.6, any value of the form $a\Delta S_T$ ($a\in\mathbb{R}$) added to the correction term does not affect the approximation.

To obtain the desired results we need to consider a dominant process $\overline{\Phi}\geq\max\{\Phi,\Phi_-\}$ because in the calculations handling the jump part of the error we end up with Φ_- which is not càdlàg and therefore is not a suitable candidate for a weight process. A prototype for $\overline{\Phi}$ is

$$\overline{\Phi}_t:=\Phi_t+\sup_{s\in[0,t]}\max\{0,-\Delta\Phi_s\},\quad t\in[0,T]. \tag{3.9}$$

For $\overline{\Phi}$ in (3.9), one has $\overline{\Phi}\in\mathrm{CL}^+([0,T])$ with $\max\{\Phi,\Phi_-\}\leq\overline{\Phi}$, and $\Phi\equiv\overline{\Phi}$ if Φ is continuous. Moreover, Proposition B.1(2) shows that $\Phi\in\mathcal{SM}_p(\mathbb{P})$ implies $\overline{\Phi}\in\mathcal{SM}_p(\mathbb{P})$.

THEOREM 3.10. *Let Assumption 3.4 hold for some $\theta\in(0,1]$ and denote $\kappa:=\frac{1-\theta}{2}$. Assume $\overline{\Phi}\geq\max\{\Phi,\Phi_-\}$ and $\overline{\Phi}\in\mathcal{SM}_2(\mathbb{P})$.*

(1) *If there is some $\alpha\in(0,2]$ such that*

$$c_{(3.10)}:=\sup_{r\in(0,1)}\left\|(\omega,t)\mapsto r^\alpha\int_{r<|z|\leq 1}\nu_t(\omega,\mathrm{d}z)\right\|_{L_\infty(\Omega\times[0,T],\mathbb{P}\otimes\lambda)}<\infty, \tag{3.10}$$

then a constant $c_{(3.11)} > 0$ *exists such that for all* $\tau \in \mathcal{T}_{\mathrm{det}}$, $\varepsilon > 0$,

$$\|E^{\mathrm{adap}}(\vartheta, \tau|\varepsilon, \kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \le c_{(3.11)} \max\{\varepsilon, \sqrt{\|\tau\|_\theta}, h(\varepsilon)\sqrt{\|\tau\|_\theta}\}, \tag{3.11}$$

where $h(\varepsilon) = \varepsilon^{1-\alpha}$ *if* $\alpha \in (1, 2]$, $h(\varepsilon) = \log^+(1/\varepsilon)$ *if* $\alpha = 1$, *and* $h(\varepsilon) = 1$ *if* $\alpha \in (0, 1)$.

(2) *If*

$$c_{(3.12)} := \sup_{r\in(0,1)} \left\| (\omega, t) \mapsto \int_{r<|z|\le 1} z\nu_t(\omega, \mathrm{d}z) \right\|_{L_\infty(\Omega\times[0,T],\mathbb{P}\otimes\lambda)} < \infty, \tag{3.12}$$

then a constant $c_{(3.13)} > 0$ *exists such that for all* $\tau \in \mathcal{T}_{\mathrm{det}}$, $\varepsilon > 0$,

$$\|E^{\mathrm{adap}}(\vartheta, \tau|\varepsilon, \kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \le c_{(3.13)} \max\{\varepsilon, \sqrt{\|\tau\|_\theta}\}. \tag{3.13}$$

We postpone the proof of Theorem 3.10 to Section 5.2.1. Minimizing the right-hand side of (3.11) and (3.13) over $\varepsilon > 0$ leads us to the following corollary.

COROLLARY 3.11. *Let Assumption* 3.4 *hold for some* $\theta \in (0, 1]$ *and denote* $\kappa := \frac{1-\theta}{2}$. *Assume* $\overline{\Phi} \ge \max\{\Phi, \Phi_-\}$ *and* $\overline{\Phi} \in \mathcal{SM}_2(\mathbb{P})$.

(1) *If* (3.10) *is satisfied for some* $\alpha \in (0, 2]$, *then there exists a constant* $c' > 0$ *independent of* τ *such that, for* $\varepsilon(\theta, \tau, \alpha) := \|\tau\|_\theta^{\frac{1}{2}(\frac{1}{\alpha}\wedge 1)}$, *one has*

$$\|E^{\mathrm{adap}}(\vartheta, \tau|\varepsilon(\theta, \tau, \alpha), \kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \le c' \begin{cases} \sqrt[2\alpha]{\|\tau\|_\theta} & \text{if } \alpha \in (1, 2], \\ \left[1 + \frac{1}{2}\log^+\left(\frac{1}{\|\tau\|_\theta}\right)\right]\sqrt{\|\tau\|_\theta} & \text{if } \alpha = 1, \\ \sqrt{\|\tau\|_\theta} & \text{if } \alpha \in (0, 1). \end{cases}$$

(2) *If* (3.12) *is satisfied, then* $\|E^{\mathrm{adap}}(\vartheta, \tau|\sqrt{\|\tau\|_\theta}, \kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \le c_{(3.13)}\sqrt{\|\tau\|_\theta}$.

REMARK 3.12.

(1) *The assumption* $j^Z_{(2.6)} < \infty$ *means that*

$$\left\| (\omega, t) \mapsto \int_{\mathbb{R}} z^2 \nu_t(\omega, \mathrm{d}z) \right\|_{L_\infty(\Omega\times[0,T],\mathbb{P}\otimes\lambda)} < \infty, \tag{3.14}$$

which implies that (3.10) *automatically holds for* $\alpha = 2$ *in our context.*

(2) *It is easy to check that condition* (3.12) *holds true if the following finite variation property is satisfied*:

$$\left\| (\omega, t) \mapsto \int_{|z|\le 1} |z|\nu_t(\omega, \mathrm{d}z) \right\|_{L_\infty(\Omega\times[0,T],\mathbb{P}\otimes\lambda)} < \infty, \tag{3.15}$$

or if the local symmetry property is satisfied: *there is an* $r_0 \in (0, 1)$ *such that the measure* $\nu_t(\omega, \cdot)$ *is symmetric on* $(-r_0, r_0)$ *for* $\mathbb{P} \otimes \lambda$-*a.e.* $(\omega, t) \in \Omega \times [0, T]$.

(3) *If* (3.10) *is satisfied for some* $\alpha \in (0, 1)$, *then* (3.15), *and hence*, (3.12) *hold true. This assertion can be deduced from applying Lemma* 5.4 (ω, t)-*wise with* $\alpha < \gamma = 1$.

3.3. *A pure-jump model with faster convergence rates*. We investigate in this part the effect of small jumps of the underlying process Z on the convergence rate of the approximation error resulting from the jump-adapted scheme.

ASSUMPTION 3.13. *In Section 2.3, we assume that*:

(a) $Z^{\mathrm{c}} \equiv 0$, $b^Z \equiv 0$, *that is*, Z *is a purely discontinuous martingale*,
(b) *the family of Lévy measures* $(\nu_t)_{t\in[0,T]}$ *does not depend on* ω,
(c) $\sigma(x) = x$, *that is*, S *is the Doléans–Dade exponential of* Z,

and assume in addition for some $\vartheta \in \mathcal{A}(S)$ *that*:

(d) $\vartheta S = (\vartheta_t S_t)_{t\in[0,T)}$ *has the semimartingale representation*

$$\vartheta_t S_t = M_t + V_t, \quad t \in [0, T),$$

where $M = (M_t)_{t\in[0,T)}$ *is an* $L_2(\mathbb{P})$*-martingale and* $V_t = \int_0^t v_u \,\mathrm{d}u$ *for a progressively measurable* v *with* $\int_0^t v_u^2(\omega)\,\mathrm{d}u < \infty$ *for* $(\omega, t) \in \Omega \times [0, T)$.

(e) *Assumption 3.4 holds for some* $\theta \in (0, 1]$, $\Phi \in \mathrm{CL}^+([0, T])$, *nondecreasing* $\Theta \in \mathrm{CL}^+([0, T])$ *with* $\Theta S \le \Phi$, *and for*

$$\Upsilon(\cdot, \mathrm{d}t) = \mathrm{d}\langle M\rangle_t + M_t^2\,\mathrm{d}t + v_t^2\,\mathrm{d}t + \left(\int_0^t v_u^2\,\mathrm{d}u\right)\mathrm{d}t, \quad t \in [0, T).$$

Let us shortly discuss conditions in Assumption 3.13. According to [25], Chapter II, Theorem 4.15, condition (b) implies that Z has independent increments. Typical examples for Z are time-inhomogeneous Lévy processes, and especially, Lévy processes. This condition is necessary for our technique exploiting the orthogonality of martingale increments which appear in the approximation error with the integrand given in (d). Condition (c) is merely a convenient condition and it can be extended to more general context with an appropriate modification for weight processes. We use here $\sigma(x) = x$ to make the proof more transparent and reduce unnecessary technicalities, and moreover, this is a classical case in application. Condition (d) is the semimartingale decomposition of ϑS, and this can be easily verified if one knows the semimartingale representation of ϑ. Condition (e) enables to obtain convergence rates for the approximation error, and this structure of Υ is discussed in Example 3.3. We will show in Proposition 4.6 below that Assumption 3.13 is fully satisfied in the exponential Lévy model when ϑ is the mean-variance hedging strategy of a European type option.

The conditions for small jump behavior of the underlying process Z are adapted respectively from (3.10) and (3.12) to the current setting as follows:

$$\sup_{r\in(0,1)} \left\| t \mapsto r^\alpha \int_{r<|z|\le 1} \nu_t(\mathrm{d}z)\right\|_{L_\infty([0,T],\lambda)} < \infty \quad \text{for some } \alpha \in (0, 2], \tag{3.16}$$

$$\sup_{r\in(0,1)} \left\| t \mapsto \int_{r<|z|\le 1} z\nu_t(\mathrm{d}z)\right\|_{L_\infty([0,T],\lambda)} < \infty. \tag{3.17}$$

It is easy to check that the condition $\|t \mapsto \int_{|z|\le 1} |z|\nu_t(\mathrm{d}z)\|_{L_\infty([0,T],\lambda)} < \infty$ implies both (3.16) (for $\alpha = 1$) and (3.17).

THEOREM 3.14. *Let Assumption 3.13 hold for some* $\theta \in (0, 1]$ *and let* $\kappa := \frac{1-\theta}{2}$. *Assume* $\overline{\Phi} \ge \max\{\Phi, \Phi_-\}$ *and* $\overline{\Phi} \in \mathcal{SM}_2(\mathbb{P})$. *If* (3.16) *is satisfied for some* $\alpha \in (0, 2]$, *then there is a* $c > 0$ *such that for all* $\tau \in \mathcal{T}_{\mathrm{det}}$, $\varepsilon > 0$,

$$\|E^{\mathrm{adap}}(\vartheta, \tau|\varepsilon, \kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})}$$
$$\le c \begin{cases} \max\{\varepsilon, \varepsilon^{1-\frac{\alpha}{2}}\sqrt{\|\tau\|_\theta}, \|\tau\|_\theta + \varepsilon^{1-\alpha}\|\tau\|_\theta^{1-\kappa(\alpha-1)}\} & \text{if } \alpha \in (1, 2], \\ \max\left\{\varepsilon, \sqrt{\varepsilon}\sqrt{\|\tau\|_\theta}, \left[1 + \log^+\left(\frac{1}{\varepsilon}\right) + \log^+\left(\frac{1}{\|\tau\|_\theta}\right)\right]\|\tau\|_\theta\right\} & \text{if } \alpha = 1, \\ \max\{\varepsilon, \sqrt{\varepsilon}\sqrt{\|\tau\|_\theta}, \|\tau\|_\theta\} & \text{if } \alpha = 1 \\ & \text{and (3.17) holds}, \\ \max\{\varepsilon, \sqrt{\varepsilon}\sqrt{\|\tau\|_\theta}, \|\tau\|_\theta\} & \text{if } \alpha \in (0, 1). \end{cases}$$

In particular, minimizing the right-hand side over $\varepsilon > 0$ yields another constant $c' > 0$ such that for all $\tau \in \mathcal{T}_{\mathrm{det}}$,

$$\|E^{\mathrm{adap}}(\vartheta, \tau|\varepsilon(\theta,\tau,\alpha),\kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \leq c' \begin{cases} \|\tau\|_\theta^{\frac{1}{\alpha}(1-\kappa(\alpha-1))} & \text{if } \alpha \in (1,2], \\ \left[1+\log^+\left(\frac{1}{\|\tau\|_\theta}\right)\right]\|\tau\|_\theta & \text{if } \alpha = 1, \\ \|\tau\|_\theta & \text{if } \alpha = 1 \text{ and (3.17) holds}, \\ \|\tau\|_\theta & \text{if } \alpha \in (0,1), \end{cases}$$

where $\varepsilon(\theta,\tau,\alpha) := \|\tau\|_\theta^{\frac{1}{\alpha}(1-\kappa(\alpha-1))}$ if $\alpha \in [1,2]$, and $\varepsilon(\theta,\tau,\alpha) := \|\tau\|_\theta$ if $\alpha \in (0,1)$.

The proof will be provided later in Section 5.2.2.

3.4. *Adapted time-nets and approximation accuracy.* We discuss in this part how to improve the approximation accuracy by using suitable time-nets.

3.4.1. *Adapted time-net.* The conclusions in Proposition 3.6, Corollary 3.11 and Theorem 3.14 assert that the errors measured in $\mathrm{bmo}_2^{\Phi}(\mathbb{P})$ or in $\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})$ are, up to multiplicative constants, upper bounded by $\|\tau\|_\theta^r$ with $r \in [\frac{1}{4}, 1]$. Assume $\tau_n \in \mathcal{T}_{\mathrm{det}}$ with $\#\tau_n = n+1$, where $n \geq 1$ represents in the context of stochastic finance the number of transactions in trading. If one uses the equidistant nets $\tau_n = (Ti/n)_{i=0}^n$, then $\|\tau_n\|_\theta = T^\theta/n^\theta$, and thus $\theta \in (0,1]$ describes the convergence rate in this situation.

In order to accelerate the convergence rate we need to employ other suitable time-nets. First, it is straightforward to check that $\|\tau_n\|_\theta \geq T^\theta/n$ for any $\tau_n \in \mathcal{T}_{\mathrm{det}}$ with $\#\tau_n = n+1$. Next, minimizing $\|\tau_n\|_\theta$ over $\tau_n \in \mathcal{T}_{\mathrm{det}}$ with $\#\tau_n = n+1$ leads us to the following adapted time-nets, which were used (at least) in [13, 14, 16, 18, 19]: For $\theta \in (0,1]$ and $n \geq 1$, the *adapted time-net* $\tau_n^\theta = (t_{i,n}^\theta)_{i=0}^n$ is defined by

$$t_{i,n}^\theta := T\big(1-(1-i/n)^{1/\theta}\big), \quad i = 0,1,\dots,n.$$

Obviously, the equidistant time-net corresponds to $\theta = 1$. By a computation, it holds that

$$T^\theta/n \leq \|\tau_n^\theta\|_\theta \leq T^\theta/(\theta n). \tag{3.18}$$

3.4.2. *Cardinality of the combined time-net.* The time-net used in Theorems 3.10 and 3.14 is $\tau \sqcup \rho(\varepsilon,\kappa)$ for $\kappa = \frac{1-\theta}{2}$. Due to the randomness, a simple way to quantify the cardinality of this combined time-net is to evaluate its expected cardinality, that is, $\mathbb{E}[\#\tau \sqcup \rho(\varepsilon,\kappa)]$ (see, e.g., [11] or [30], equation (10) with $\beta = 0$). Thus, we provide in the next result an estimate for certain moments of the cardinality. Since we aim to apply Proposition 2.5(3) later, changes of the underlying measure are also taken into account.

PROPOSITION 3.15. *Let $q \in [1,2]$, $r \in [2,\infty]$ with $\frac{q}{2} + \frac{1}{r} = 1$. Assume a probability measure $\mathbb{Q}$ absolutely continuous with respect to $\mathbb{P}$ with $\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P} \in L_r(\mathbb{P})$. If (3.10) holds for some $\alpha \in (0,2]$, then for any $\kappa \in [0, \frac{1}{2})$ and $(\varepsilon_n)_{n\geq 1} \subset (0,\infty)$ with $\inf_{n\geq 1} \sqrt[\alpha]{n}\varepsilon_n > 0$, there exists a constant $c_{(3.19)} \geq 1$ such that for any $n \geq 1$, any $\tau_n \in \mathcal{T}_{\mathrm{det}}$ with $\#\tau_n = n+1$ one has*

$$\|\#\tau_n \sqcup \rho(\varepsilon_n,\kappa)\|_{L_q(\mathbb{Q})} \sim_{c_{(3.19)}} n. \tag{3.19}$$

Plugging the adapted time-nets τ_n^θ into previous results, we derive the following.

THEOREM 3.16. *Let Assumption 3.4 hold true for some $\theta \in (0,1]$ and let $\kappa := \frac{1-\theta}{2}$.*

(1) *One has* $\sup_{n\geq 1} \sqrt{n}\|E^{\mathrm{Rm}}(\vartheta, \tau_n^\theta)\|_{\mathrm{bmo}_2^\Phi(\mathbb{P})} < \infty$.

(2) *Assume* $\overline{\Phi} \geq \max\{\Phi, \Phi_-\}$ *and* $\overline{\Phi} \in \mathcal{SM}_2(\mathbb{P})$.

  (a) *If* (3.10) *holds for some* $\alpha \in (0,2]$, *then*

$$\sup_{n\geq 1} R(n)\|E^{\mathrm{adap}}(\vartheta, \tau_n^\theta|\varepsilon_n, \kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} < \infty,$$

where $R(n) = 1/\varepsilon_n = \sqrt[2\alpha]{n}$ *if* $\alpha \in (1,2]$, $R(n) = \sqrt{n}/(1+\log n)$ *and* $\varepsilon_n = \sqrt{1/n}$ *if* $\alpha = 1$, *and* $R(n) = 1/\varepsilon_n = \sqrt{n}$ *if* $\alpha \in (0,1)$.

  (b) *If* (3.12) *holds, then for* $R(n) = 1/\varepsilon_n = \sqrt{n}$ *one has*

$$\sup_{n\geq 1} R(n)\|E^{\mathrm{adap}}(\vartheta, \tau_n^\theta|\varepsilon_n, \kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} < \infty.$$

(3) *Assume Assumption* 3.13 *and* $\overline{\Phi} \geq \max\{\Phi, \Phi_-\}$ *with* $\overline{\Phi} \in \mathcal{SM}_2(\mathbb{P})$. *If* (3.16) *holds for some* $\alpha \in (0,2]$, *then*

$$\sup_{n\geq 1} R(n)\|E^{\mathrm{adap}}(\vartheta, \tau_n^\theta|\varepsilon_n, \kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} < \infty,$$

where

$$\begin{cases} R(n) = 1/\varepsilon_n = n^{\frac{1}{\alpha}(1-\kappa(\alpha-1))} & \text{if } \alpha \in (1,2], \\ R(n) = n/(1+\log n),\ \varepsilon_n = 1/n & \text{if } \alpha = 1, \\ R(n) = 1/\varepsilon_n = n & \text{if } \alpha = 1 \text{ and (3.17) holds}, \\ R(n) = 1/\varepsilon_n = n & \text{if } \alpha \in (0,1). \end{cases}$$

(4) *If in addition* $\overline{\Phi} \in \mathcal{SM}_p(\mathbb{P})$ *for some* $p > 2$, *then the conclusions in Items* (2)–(3) *hold case-wise for the* $\mathrm{BMO}_p^{\overline{\Phi}}(\mathbb{P})$*-norm and, consequently, for the* $S_p(\mathbb{P})$*-norm, in place of the* $\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})$*-norm.*

(5) *If in addition* $\mathbb{Q} \in \mathcal{RH}_s(\mathbb{P})$ *for some* $s \in (1,\infty)$ *and* $\overline{\Phi} \in \mathcal{SM}_2(\mathbb{Q})$, *then the conclusions in Items* (2)–(3) *hold case-wise for the* $\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{Q})$*-norm in place of the* $\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})$*-norm.*

PROOF. Item (1) (resp. (2), (3)) follows from Proposition 3.6 (resp. Theorem 3.10, Theorem 3.14) and (3.18). Items (4)–(5) are due to Proposition 2.5. □

REMARK 3.17. *In Theorem* 3.16, *the convergence rate of the jump-adapted approximation errors is* $R(n)^{-1}$ *as* $n \to \infty$. *However, it is necessary to quantify the convergence rate in terms of expected cardinality of the used discretization times, which is* $\tau_n^\theta \sqcup \rho(\varepsilon_n, \kappa)$. *It turns out that the rates remain unchanged as shown below.*

(1) *For Items* (2a), (3) *and* (4), *applying Proposition* 3.15 *with* $q = 2$, $r = \infty$ *and* $\mathbb{Q} = \mathbb{P}$ *we find that* $\|\#\tau_n \sqcup \rho(\varepsilon_n, \kappa)\|_{L_2(\mathbb{P})} \sim_{c_{(3.19)}} n$. *Consequently,* $\|\#\tau_n \sqcup \rho(\varepsilon_n, \kappa)\|_{L_u(\mathbb{P})} \sim_{c_{(3.19)}} n$ *for any* $u \in [1,2]$. *Hence, there exists a constant* $c \geq 1$ *not depending on* n *such that*

$$R(n) \sim_c R(\|\#\tau_n^\theta \sqcup \rho(\varepsilon_n, \kappa)\|_{L_u(\mathbb{P})}), \quad \forall u \in [1,2] \tag{3.20}$$

for R *and* $(\varepsilon_n)_{n\geq 1}$ *given case-wise as above.*

(2) *For Item* (2b), *we apply Proposition* 3.15 *with* $q = 2$, $r = \infty$, $\mathbb{Q} = \mathbb{P}$ *and* $\alpha = 2$ (*with keeping* (3.14) *in mind*) *to obtain* (3.20).

(3) *For Item* (5), *if* $s \in [2,\infty)$, *then we choose* $r = s$, $q = \tilde{s} := 2(1-\frac{1}{s}) \in [1,2)$ *in Proposition* 3.15 *to get a constant* $c' \geq 1$ *not depending on* n *such that*

$$R(n) \sim_{c'} R(\|\#\tau_n^\theta \sqcup \rho(\varepsilon_n, \kappa)\|_{L_{\tilde{s}}(\mathbb{Q})}).$$

3.5. *A lower bound for jump-adapted approximation errors.* This subsection provides a situation where one achieves the sharp convergence rates in Theorem 3.16. Namely, we illustrate in Example 3.19 below that the obtained rates in Theorem 3.16(2) and (3) are sharp if $\alpha \in (0,2]\backslash\{1\}$, and are sharp up to a log-factor if $\alpha = 1$. We begin with the following result.

PROPOSITION 3.18. *Let $\kappa := \frac{1-\theta}{2} \in [0, \frac{1}{2})$. Under the setting of Section 2.3, we assume a process $0 < \overline{\Phi} \in \mathcal{SM}_2(\mathbb{P})$, constants $\varepsilon_0 > 0$ and $c_{(3.21)} > 0$ such that for any $\varepsilon \in (0, \varepsilon_0)$ and $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\det}$ with $\|\tau\|_1 \le T/2$ there exist an $i_* \in \{1, \dots, n\}$, real numbers $0 < r_* < \hat{r}_* \le T^\kappa$ and a stopping time $\rho_* \colon \Omega \to [0, T)$ such that*

$$\begin{aligned}\mathbb{P}(E_{(3.21)}) := \mathbb{P}\bigg(&\{\rho_* \in (t_{i_*-1}, t_{i_*}]\} \cap \bigg\{\frac{\sigma(S_{\rho_*-})}{\overline{\Phi}_{\rho_*}}|\vartheta_{\rho_*-} - \vartheta_{t_{i_*-1}-}| \ge \frac{c_{(3.21)}}{r_*}\bigg\} \\ &\cap \big\{|\Delta Z_{\rho_*}| \in (\varepsilon r_*, \varepsilon \hat{r}_*)\big\} \cap \big\{\hat{r}_* \le (T-\rho_*)^\kappa\big\}\bigg) > 0.\end{aligned} \tag{3.21}$$

Then, a constant $c_{(3.22)} > 0$ exists such that for any $\varepsilon \in (0, \varepsilon_0)$ and $\tau \in \mathcal{T}_{\det}$ with $\|\tau\|_1 \le T/2$,

$$\big\|E^{\mathrm{adap}}(\vartheta, \tau|\varepsilon, \kappa)\big\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \ge c_{(3.22)}\varepsilon. \tag{3.22}$$

PROOF. We employ the notation from the proof of Theorem 3.10. One first notices that

$$\big\{|\Delta Z_{\rho_*}| \in (\varepsilon r_*, \varepsilon \hat{r}_*)\big\} \cap \big\{\hat{r}_* \le (T-\rho_*)^\kappa\big\} = \big\{\big|\Delta Z_{\rho_*}^{\varepsilon,1}\big| \in (\varepsilon r_*, \varepsilon \hat{r}_*)\big\} \cap \big\{\hat{r}_* \le (T-\rho_*)^\kappa\big\},$$

and that $\Delta E^{\mathrm{adap}}(\vartheta, \tau|\varepsilon, \kappa) = \Delta E^{\mathrm{S}}(\vartheta, \tau|\varepsilon, \kappa)$ by (5.10). We then apply Lemma 2.3 to obtain a constant $c_{\overline{\Phi}} > 0$ depending on $\|\overline{\Phi}\|_{\mathcal{SM}_2(\mathbb{P})}$ such that for any $\varepsilon \in (0, \varepsilon_0)$ and $\tau \in \mathcal{T}_{\det}$, a.s.,

$$\begin{aligned}\big\|E^{\mathrm{adap}}(\vartheta, \tau|\varepsilon, \kappa)\big\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} &\ge c_{\overline{\Phi}}\big|\Delta E^{\mathrm{adap}}(\vartheta, \tau|\varepsilon, \kappa)\big|_{\overline{\Phi}} = c_{\overline{\Phi}}\big|\Delta E^{\mathrm{S}}(\vartheta, \tau|\varepsilon, \kappa)\big|_{\overline{\Phi}} \\ &= c_{\overline{\Phi}} \sup_{t\in[0,T)} \sum_{i=1}^n \bigg[\frac{\sigma(S_{t-})}{\overline{\Phi}_t}|\vartheta_{t-} - \vartheta_{t_{i-1}-}|\big|\Delta Z_t^{\varepsilon,1}\big|\bigg]\mathbb{1}_{(t_{i-1},t_i]}(t) \\ &\ge c_{\overline{\Phi}} \sup_{t\in(t_{i_*-1},t_{i_*}]\cap(0,T)} \bigg[\frac{\sigma(S_{t-})}{\overline{\Phi}_t}|\vartheta_{t-} - \vartheta_{t_{i_*-1}-}|\big|\Delta Z_t^{\varepsilon,1}\big|\bigg] \\ &\ge c_{\overline{\Phi}} \frac{\sigma(S_{\rho_*-})}{\overline{\Phi}_{\rho_*}}|\vartheta_{\rho_*-} - \vartheta_{t_{i_*-1}-}|\big|\Delta Z_{\rho_*}^{\varepsilon,1}\big|\mathbb{1}_{(t_{i_*-1},t_{i_*}]}(\rho_*).\end{aligned}$$

Thus, a.s.,

$$\big\|E^{\mathrm{adap}}(\vartheta, \tau|\varepsilon, \kappa)\big\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \ge c_{\overline{\Phi}} \frac{c_{(3.21)}}{r_*}\varepsilon r_* \mathbb{1}_{E_{(3.21)}} = c_{\overline{\Phi}} c_{(3.21)} \varepsilon \mathbb{1}_{E_{(3.21)}}.$$

Since $\mathbb{P}(E_{(3.21)}) > 0$ by assumption, we get (3.22) with $c_{(3.22)} = c_{\overline{\Phi}} c_{(3.21)}$. □

Remark that condition (3.21) describes the high fluctuation of ϑ along any time net $\tau \in \mathcal{T}_{\det}$ which is captured by the stopping time ρ_*.

EXAMPLE 3.19. *Set $T = 1$. Let X be a square integrable Lévy process given by*

$$X_t = \hat{\gamma} t + \int_0^t \int_{\mathbb{R}_0} x \widetilde{N}(\mathrm{d}s, \mathrm{d}x), \quad t \in [0, 1],$$

for some $\hat{\gamma} \in \mathbb{R}$, where N is the Poisson random measure of X with Lévy measure ν, and $\widetilde{N}(\mathrm{d}s, \mathrm{d}x) := N(\mathrm{d}s, \mathrm{d}x) - \mathrm{d}s\nu(\mathrm{d}x)$ (see, e.g., [1], Theorem 2.4.16). Set $S := \mathrm{e}^X$ and assume that:

- $\int_{|x|>1} \mathrm{e}^{2x}\nu(\mathrm{d}x) < \infty$, $(0,1) \subset \operatorname{supp}(\nu)$, *and there is a constant* $c_{(3.23)} \in (0,1]$ *with*

$$(-\infty, \ln c_{(3.23)}) \cap \operatorname{supp}(\nu) = \varnothing, \tag{3.23}$$

  where $\operatorname{supp}(\nu)$ *denotes the support of* ν.
- $\vartheta_t = \phi(t, S_t)$, *where* $\phi\colon [0,1) \times (0,\infty)$ *is jointly continuous, such that* $\mathbb{E}\int_0^T \vartheta_t^2 S_t^2\,\mathrm{d}t < \infty$ *and that constants* $T_0 \in (0,1)$, $K > 0$ *exist such that* $y \mapsto \phi(t,y)$ *is continuously differentiable for all* $t \in (0, T_0]$ *and*

$$\inf_{(t,y)\in(0,T_0]\times[1,\mathrm{e}^{(5K+11)/2}]} |\partial_y \phi(t,y)| \geq c_{(3.24)} > 0. \tag{3.24}$$

  Condition (3.24) *is analogous to and slightly weaker than that in* [30], *Proposition* 4. *A simple example for* (3.24) *is* $\phi(t,y) = y$, *and then the approximated integral becomes*

$$\int_0^T S_{t-}\,\mathrm{d}S_t.$$

- *The weight processes are* $\Theta(\eta)_t := \sup_{u\leq t}(S_u^{\eta-1})$ *and* $\Phi(\eta) := \Theta(\eta)S$ *for some* $\eta \in [0,1]$.

Then the assumptions of Proposition 3.18 *are satisfied for* $\sigma(x) = x$, $Z := \mathcal{L}(S)$ *the stochastic logarithm of* S, *and for*

$$\begin{aligned} &\overline{\Phi} := \frac{\Phi(\eta)}{c_{(3.23)}}, \qquad \varepsilon_0 := 1, \qquad c_{(3.21)} := \frac{1}{2^\kappa 12} c_{(3.23)} c_{(3.24)} \mathrm{e}^{-(1-\eta)(\frac{K}{2}+1)-1} K, \\ &i_* := 1, \qquad \hat{r}_* = 3r_* := 2^{-\kappa}, \qquad \rho_* := \rho_{K_{\tilde{\varepsilon}}} \quad \text{given in (5.30) with } \tilde{\varepsilon} = \ln(1+\varepsilon r_*). \end{aligned} \tag{3.25}$$

We will verify this assertion later in Section 5.2.4.

REMARK 3.20. *Let us give some comments on Example* 3.19.

(1) *To validate Assumption* 3.13, *we need to additionally assume the following conditions, which are feasible because of Proposition* 4.6:

- S *is an* $L_2(\mathbb{P})$*-martingale.*
- *There are constants* $\theta \in (0,1]$ *and* $c > 0$ *such that*

$$|\phi(t, S_t)| \leq c(1-t)^{\frac{\theta-1}{2}} \Theta(\eta)_t \quad a.s., \forall t \in [0,1).$$

We may even assume that ϑS *is an* $L_2(\mathbb{P})$*-martingale to simplify Assumption* 3.13.

(2) *The selection* $\overline{\Phi} := \Phi(\eta)/c_{(3.23)}$ *in* (3.25) *satisfies the condition for the weight process in Theorems* 3.10, 3.14 *and* 3.16. *Indeed, since* $\mathrm{e}^{\Delta X} \geq c_{(3.23)}$ *and* $\Theta(\eta)$ *is nondecreasing, one has*

$$\Phi(\eta)_{t-} = \Theta(\eta)_{t-} S_{t-} = \Theta(\eta)_{t-} S_t \mathrm{e}^{-\Delta X_t} \leq \Theta(\eta)_t (S_t / c_{(3.23)}) = \Phi(\eta)_t / c_{(3.23)}.$$

Since $c_{(3.23)} \in (0,1]$, *it implies that* $\overline{\Phi} \geq \max\{\Phi(\eta), \Phi(\eta)_-\}$.

(3) *When the assumptions of Theorem* 3.16(2) *or* (3) *are fulfilled, then the convergence rate of the approximation error is given by* (ε_n) (*up to a* log*-factor for the case* $\alpha = 1$). *Then we apply Proposition* 3.18 *with the choice* $\varepsilon = \varepsilon_n$ *to find that, for sufficiently large* n, *the obtained convergence rates in Theorem* 3.16(2) *or* (3) *are optimal if* $\alpha \in (0,2]\backslash\{1\}$, *and are optimal up to a* log*-factor if* $\alpha = 1$.

4. Applications to exponential Lévy models. We provide examples for Assumptions 3.4 and 3.13 in the exponential Lévy setting so that the main results can be applied. As an important step to obtain them, we establish in Proposition 4.2 an explicit form for the mean-variance hedging strategy of a general European type option, and this formula might also have an independent interest.

4.1. *Lévy process.* Let $X = (X_t)_{t\in[0,T]}$ be a one-dimensional Lévy process on $(\Omega, \mathcal{F}, \mathbb{P})$, that is, $X_0 = 0$, X has independent and stationary increments and X has càdlàg paths. Let $\mathbb{F}^X = (\mathcal{F}_t^X)_{t\in[0,T]}$ be the augmented natural filtration of X, and assume that $\mathcal{F} = \mathcal{F}_T^X$. According to the Lévy–Khintchine formula (see, e.g., [31], Theorem 8.1), there is a *characteristic triplet* (γ, σ, ν), where $\gamma \in \mathbb{R}$, coefficient of Brownian component $\sigma \geq 0$, Lévy measure $\nu\colon \mathcal{B}(\mathbb{R}) \to [0, \infty]$ (i.e., $\nu(\{0\}) := 0$ and $\int_{\mathbb{R}}(x^2 \wedge 1)\nu(\mathrm{d}x) < \infty$), such that the *characteristic exponent* ψ of X defined by $\mathbb{E}\mathrm{e}^{\mathrm{i}uX_t} = \mathrm{e}^{-t\psi(u)}$ is of the form

$$\psi(u) = -\mathrm{i}\gamma u + \frac{\sigma^2 u^2}{2} - \int_{\mathbb{R}} (\mathrm{e}^{\mathrm{i}ux} - 1 - \mathrm{i}ux\mathbb{1}_{\{|x|\leq 1\}})\nu(\mathrm{d}x), \quad u \in \mathbb{R}.$$

4.2. *Mean-variance hedging* (*MVH*). Assume that the underlying price process is modelled by the exponential $S = \mathrm{e}^X$. Since models with jumps correspond to incomplete markets in general, there is no optimal hedging strategy which replicates a payoff at maturity and eliminates risks completely. This leads to consider certain strategies that minimize some types of risk. Here, we use quadratic hedging which is a common approach, see [33]. To simplify the quadratic hedging problem, we consider the martingale market. Applications of results in Section 3 for Lévy markets under the semimartingale setting are studied in [36].

ASSUMPTION 4.1. *$S = \mathrm{e}^X$ is an $L_2(\mathbb{P})$-martingale and is not a.s. constant.*

Under Assumption 4.1, any $\xi \in L_2(\mathbb{P})$ admits the *Galtchouk–Kunita–Watanabe (GKW) decomposition*

$$\xi = \mathbb{E}\xi + \int_0^T \theta_t^\xi \,\mathrm{d}S_t + L_T^\xi, \tag{4.1}$$

where θ^ξ is predictable with $\mathbb{E}\int_0^T |\theta_t^\xi|^2 S_{t-}^2 \,\mathrm{d}t < \infty$, $L^\xi = (L_t^\xi)_{t\in[0,T]}$ is an $L_2(\mathbb{P})$-martingale with zero mean and satisfies $\langle S, L^\xi\rangle = 0$. The integrand θ^ξ is called the *MVH strategy* for ξ, which is unique in $L_2(\Omega \times [0, T], \mathbb{P} \otimes \lambda)$. The reader is referred to [33] for further discussion.

Our aim is to apply the approximation results obtained in Section 3 for the stochastic integral term in (4.1), which can be interpreted in mathematical finance as the hedgeable part of ξ. To do that, one of our main tasks is to find a representation of θ^ξ which is convenient for verifying assumptions in Section 3. This issue is handled in Section 4.3 in which we focus on the European type options $\xi = g(S_T)$.

4.3. *Explicit MVH strategy.* In the literature, there are several methods to determine an explicit form for the MVH strategy of a European type option $g(S_T)$. Let us mention some typical approaches for which the martingale representation of $g(S_T)$ plays the key role. A classical method is by using directly Itô's formula (e.g., [7, 24]) which requires a certain smoothness of $(t, y) \mapsto \mathbb{E}g(yS_{T-t})$. Another idea is based on Fourier analysis to separate the payoff function g and the underlying process S (e.g., [4, 23, 34]). To do that, some regularity for g and S is assumed. As a third method, one can use Malliavin calculus to determine the MVH strategy (e.g., [2]), however the payoff $g(S_T)$ is assumed to be differentiable in the Malliavin sense so that the Clark–Ocone formula is applicable.

To the best of our knowledge, Proposition 4.2 below is new and it provides an explicit formula for the MVH strategy of $g(S_T)$ without requiring any regularity from the payoff function g nor any specific structure of the underlying process S. Recall that σ and ν are the coefficient of the Brownian component and the Lévy measure of X respectively.

PROPOSITION 4.2. *Assume Assumption 4.1. For a Borel function $g\colon \mathbb{R}_+ \to \mathbb{R}$ with $g(S_T) \in L_2(\mathbb{P})$, there exists a $\vartheta^g \in \mathrm{CL}([0, T))$ such that the following assertions hold*:

(1) *ϑ^g_- is a MVH strategy of $g(S_T)$;*
(2) *$\vartheta^g S$ is an $L_2(\mathbb{P})$-martingale and $\vartheta^g_t = \vartheta^g_{t-}$ a.s. for each $t \in [0,T)$;*
(3) *For any $t \in (0,T)$, one has, a.s.,*

$$\vartheta^g_t = \frac{1}{c^2_{(4.2)}}\left[\sigma^2 \partial_y G(t,S_t) + \int_{\mathbb{R}} \frac{G(t,\mathrm{e}^x S_t) - G(t,S_t)}{S_t}(\mathrm{e}^x - 1)\nu(\mathrm{d}x)\right], \tag{4.2}$$

where $c^2_{(4.2)} := \sigma^2 + \int_{\mathbb{R}}(\mathrm{e}^x - 1)^2\nu(\mathrm{d}x)$ and $G(t,\cdot)\colon \mathbb{R}_+ \to \mathbb{R}$ is as follows:

(a) *If $\sigma > 0$, then we choose $G(t,y) := \mathbb{E}g(yS_{T-t})$;*

(b) *If $\sigma = 0$, then we choose $G(t,\cdot)$ such that it is Borel measurable and $G(t,S_t) = \mathbb{E}^{\mathcal{F}_t}[g(S_T)]$ a.s., and we set $\partial_y G(t,\cdot) := 0$ by convention.*

PROOF. The proof is given in the Supplementary Material [37], Section E, by exploiting Malliavin calculus. □

Assumption 4.1 ensures that $c_{(4.2)} \in (0,\infty)$. For the case (3a), the function $G(t,\cdot)$ has derivatives of all orders on $\mathbb{R}_+$ due to the presence of the Gaussian component of X, see [17], Theorem 9.13. Formula (4.2) was also established in [7], Section 4,[5] and in [34], Proposition 7,[6] under some extra conditions for g and S. A similar formula of (4.2) in a general setting can be found in [24], Theorem 2.4.[7]

4.4. *Growth of the MVH strategy and weight regularity.* We investigate the growth in time of ϑ^g obtained in (4.2) pathwise and relatively to a weight process for Hölder continuous or bounded functions g. This growth property is examined in connection to the small jump behavior of the underlying Lévy process.

DEFINITION 4.3.

(1) (*Hölder spaces*) *Let $\emptyset \neq U \subseteq \mathbb{R}$ be an open interval and let $\eta \in [0,1]$. For a Borel function $f\colon U \to \mathbb{R}$, we define*

$$|f|_{C^{0,\eta}(U)} := \inf\{c \in [0,\infty) : |f(x) - f(y)| \le c|x-y|^\eta \text{ for all } x,y \in U, x \neq y\},$$

and let $f \in C^{0,\eta}(U)$ if $|f|_{C^{0,\eta}(U)} < \infty$. It is clear that, on U, the space $C^{0,1}(U)$ consists of all Lipschitz functions, $C^{0,\eta}(U)$ contains all η-Hölder continuous functions for $\eta \in (0,1)$, and $C^{0,0}(U)$ consists of all bounded (not necessarily continuous) Borel functions.

(2) (*α-stable-like Lévy measures*) *For a Lévy measure ν and for some $\alpha \in (0,2]$, we let*

$$\nu \in \mathscr{US}(\alpha) \quad \Leftrightarrow \quad \sup_{r\in(0,1)} r^\alpha \int_{r<|x|\le 1} \nu(\mathrm{d}x) < \infty,$$

and let $\nu \in \mathscr{S}(\alpha)$ for some $\alpha \in (0,2)$ if $\nu = \nu_1 + \nu_2$, where ν_1, ν_2 are Lévy measures with

$$\nu_1(\mathrm{d}x) = \frac{k(x)}{|x|^{\alpha+1}}\mathbb{1}_{\{x\neq 0\}}\,\mathrm{d}x \quad \text{and} \quad \nu_2 \in \mathscr{US}(\alpha),$$

where $0 < \liminf_{x\to 0} k(x) \le \limsup_{x\to 0} k(x) < \infty$, and the function $x \mapsto k(x)/|x|^\alpha$ is nondecreasing on $(-\infty,0)$ and is nonincreasing on $(0,\infty)$.

---

[5] [7], Section 4, equation (4.1), assumes either that the function g is Lipschitz or that g has certain growth and small jumps of X behave like α-stable processes.

[6] [34], Proposition 7, assumes that the payoff function g, after multiplying with an exponential damping factor, is of finite variation and belongs to $L_1(\mathbb{R})$. In addition, the characteristic exponent of X is required to satisfy a certain integrability.

[7] [24], Theorem 2.4, establishes a similar representation in a more general setting than the Lévy setting. However, one needs to assume that $(t,y) \mapsto P_t g(y)$ is a $C^{1,2}$-function so that Itô's formula is applicable. Here P_t is the semigroup associated with X.

REMARK 4.4 (see also [36], Lemma B.1). *Let ν be a Lévy measure and $\alpha \in (0,2)$.*

(1) *If $\nu \in \mathscr{S}(\alpha)$, then $\nu \in \mathscr{U}\mathscr{S}(\alpha)$ and α is equal to the Blumenthal–Getoor index of ν, that is*, $\alpha = \inf\{q \in [0,2] : \int_{|x|\le 1} |x|^q \nu(\mathrm{d}x) < \infty\}$, *see* [3].

(2) *One has $\nu \in \mathscr{S}(\alpha)$ if ν has a density $p(x) := \nu(\mathrm{d}x)/\mathrm{d}x$ satisfying*

$$0 < \liminf_{|x|\to 0} |x|^{1+\alpha} p(x) \le \limsup_{|x|\to 0} |x|^{1+\alpha} p(x) < \infty.$$

EXAMPLE 4.5. *We provide typical examples in mathematical finance using $C^{0,\eta}(U)$-payoff functions and α-stable-like processes.*

(1) *Let $K > 0$. The binary payoff $g_0(x) := \mathbb{1}_{(K,\infty)}(x)$ belongs to $C^{0,0}(\mathbb{R}_+)$ obviously, the call payoff $g_1(x) := (x-K)^+$ is contained in $C^{0,1}(\mathbb{R}_+)$, and for $\eta \in (0,1)$, the powered call payoff* (*see, e.g.*, [23]) *$g_\eta(x) := ((x-K)^+)^\eta$ belongs to $C^{0,\eta}(\mathbb{R}_+)$.*

(2) *The CGMY process* (*see* [32], *Section* 5.3.9) *with parameters $C, G, M > 0$ and $Y \in (0,2)$ has the Lévy measure*

$$\nu_{\mathrm{CGMY}}(\mathrm{d}x) = C \frac{\mathrm{e}^{Gx}\mathbb{1}_{\{x<0\}} + \mathrm{e}^{-Mx}\mathbb{1}_{\{x>0\}}}{|x|^{1+Y}} \mathbb{1}_{\{x\neq 0\}}\,\mathrm{d}x$$

which belongs to $\mathscr{S}(Y)$ due to Remark 4.4(2).

The normal inverse Gaussian (*NIG*) *process* (*see* [32], *Section* 5.3.8) *has the Lévy density $p_{\mathrm{NIG}}(x) := \nu_{\mathrm{NIG}}(\mathrm{d}x)/\mathrm{d}x$ that satisfies*

$$0 < \liminf_{|x|\to 0} x^2 p_{\mathrm{NIG}}(x) \le \limsup_{|x|\to 0} x^2 p_{\mathrm{NIG}}(x) < \infty.$$

Hence, *Remark* 4.4(2) *verifies that $\nu_{\mathrm{NIG}} \in \mathscr{S}(1)$.*

Before stating the main result of this part, let us introduce the relevant weight processes. For $\eta \in [0,1]$, define processes $\Theta(\eta), \Phi(\eta) \in \mathrm{CL}^+([0,T])$ by setting

$$\Theta(\eta)_t := \sup_{u\in[0,t]} \left(S_u^{\eta-1}\right) \quad \text{and} \quad \Phi(\eta)_t := \Theta(\eta)_t S_t. \tag{4.3}$$

Proposition 4.6 below verifies Assumptions 3.4 and 3.13 in the exponential Lévy setting, and its proof is given later in Section 5.3.

PROPOSITION 4.6. *Assume Assumption* 4.1. *Let $\eta \in [0,1]$ and $g \in C^{0,\eta}(\mathbb{R}_+)$.*

(1) (*Weight regularity*) *One has $\Phi(\eta) \in \mathcal{SM}_2(\mathbb{P})$.*

(2) (*MVH strategy growth*) *There is a constant $c_{(4.4)} > 0$ such that, for ϑ^g given in* (4.2),

$$\left|\vartheta_t^g\right| \le c_{(4.4)} U(t) S_t^{\eta-1} \quad \text{a.s.}, \forall t \in [0,T), \tag{4.4}$$

where the function $U(t)$ is provided in Table 1.

(3) *Denote $M := \vartheta^g S$. Then Assumption* 3.4 *holds true for*

$$\vartheta = \vartheta^g, \qquad \Upsilon(\cdot, \mathrm{d}t) = \mathrm{d}\langle M\rangle_t + M_t^2\,\mathrm{d}t, \qquad \Theta = \Theta(\eta), \qquad \Phi = \Phi(\eta)$$

and for θ provided in Table 1 *accordingly. In particular, if $\sigma = 0$* (*i.e., X does not have a Brownian component*), *then Assumption* 3.13 *is satisfied.*

Eventually, let us turn to the approximation problem in the exponential Lévy setting. Although results in Section 3 are stated in terms of the characteristic of Z (the integrator of the SDE (5.32)), the result below are formulated for the characteristic of the log price process X which is slightly more convenient to verify in practice. Based on the relation between X and Z in Appendix A, we can easily translate conditions imposed on X to Z and vice versa.

TABLE 1
Conclusions for $U(t)$ and θ

	σ and η	Small jump condition for X	Function $U(t)$	Values of θ
A1	$\sigma > 0$ $\eta \in [0, 1]$		$U(t) = (T - t)^{\frac{\eta-1}{2}}$	$\theta = 1$ if $\eta = 1$, $\forall\theta \in (0, \eta)$ if $\eta \in (0, 1)$
A2	$\sigma = 0$ $\eta \in [0, 1]$	$\int_{\lvert x\rvert\le 1} \lvert x\rvert^{1+\eta}\nu(\mathrm{d}x) < \infty$	$U(t) = 1$	$\theta = 1$
A3	$\sigma = 0$ $\eta \in [0, 1)$	$\nu \in \mathscr{S}(\alpha)$ for some $\alpha \in [1 + \eta, 2)$	$U(t) = (T - t)^{\frac{1+\eta}{\alpha}-1}$ if $\alpha \in (1 + \eta, 2)$, $U(t) = \max\{1, \log \frac{1}{T-t}\}$ if $\alpha = 1 + \eta$	$\forall\theta \in \big(0, \frac{2(1+\eta)}{\alpha} - 1\big)$

TABLE 2
Convergence rate $R(n)^{-1}$ and jump size threshold ε_n

	Interplay between g and X	Values of θ	$R(n)$ and ε_n
B1	$\sigma = 0$ and $\nu \in \mathscr{U}\mathscr{S}(\alpha)$ for some $(\eta, \alpha) \in ([0, 1) \times (0, 1+\eta))$ $\cup(\{1\} \times (0, 2])$	$\theta = 1$	$R(n) = 1/\varepsilon_n = \sqrt[\alpha]{n}$ if $\alpha \in (1, 2]$, $R(n) = n/(1 + \log n)$, $\varepsilon_n = 1/n$ if $\alpha = 1$, $R(n) = 1/\varepsilon_n = n$ if $\alpha \in (0, 1)$
B2	$\sigma = 0$ and $\nu \in \mathscr{S}(\alpha)$ for some $(\eta, \alpha) \in [0, 1) \times [1 + \eta, 2)$	$\forall\theta \in \big(0, \frac{2(1+\eta)}{\alpha} - 1\big)$	$R(n) = 1/\varepsilon_n = n^{\frac{1}{\alpha}(1-\frac{1}{2}(1-\theta)(\alpha-1))}$ if $(\eta, \alpha) \neq (0, 1)$, $R(n) = n/(1 + \log n)$, $\varepsilon_n = 1/n$ if $(\eta, \alpha) = (0, 1)$
B3	$\sigma > 0$ and $\eta = 1$	$\theta = 1$	$R(n) = 1/\varepsilon_n = \sqrt{n}$
B4	$\sigma > 0$, $\eta \in (0, 1)$ and $\nu \in \mathscr{U}\mathscr{S}(\alpha)$ for some $\alpha \in (0, 2]$	$\forall\theta \in (0, \eta)$	$R(n) = 1/\varepsilon_n = \sqrt{n}$ if $\alpha \in (0, \frac{3-\theta}{2-\theta}]$, $R(n) = 1/\varepsilon_n = n^{\frac{1}{\alpha}(1-\frac{1}{2}(1-\theta)(\alpha-1))}$ if $\alpha \in (\frac{3-\theta}{2-\theta}, 2]$

COROLLARY 4.7. *Assume Assumption 4.1 and let $\eta \in [0, 1]$, $g \in C^{0,\eta}(\mathbb{R}_+)$.*

(1) *For ϑ^g given in (4.2), $\Phi(\eta)$ in (4.3) and $\overline{\Phi}(\eta)$ in (3.9), one has*

$$\sup_{n\ge 1} R(n)\big\| E^{\mathrm{adap}}\big(\vartheta^g, \tau_n^\theta \,\big|\, \varepsilon_n, \kappa\big)\big\|_{\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P})} < \infty, \tag{4.5}$$

where $\kappa = \frac{1-\theta}{2}$, and θ, $R(n)$, ε_n are provided in Table 2.

(2) *If $\int_{|x|>1} \mathrm{e}^{px}\nu(\mathrm{d}x) < \infty$ for some $p > 2$, then (4.5) holds for the $\mathrm{BMO}_p^{\overline{\Phi}(\eta)}(\mathbb{P})$-norm and, consequently, for the $S_p(\mathbb{P})$-norm, in place of the $\mathrm{BMO}_2^{\overline{\Phi}(\eta)}(\mathbb{P})$-norm.*

According to Proposition 3.15 with $q = 2$, $r = \infty$, $\mathbb{Q} = \mathbb{P}$, the parameter n in (4.5) is comparable to the $L_2(\mathbb{P})$-norm of the cardinality of the combined time-nets $\tau_n^\theta \sqcup \rho(\varepsilon_n, \kappa)$ used in the approximation.

5. Proofs of results in Section 3 and Section 4.4.

5.1. *Proofs of results in Section 3.1.* We need the following auxiliary result.

LEMMA 5.1. *There are constants* $c_{(5.1)}, c_{(5.2)} > 0$ *such that for any* $0 \le a < b \le T$, *a.s.*,

$$\mathbb{E}^{\mathcal{F}_a}\left[\int_a^b \sigma(S_t)^2\,\mathrm{d}t\right] \le c_{(5.1)}^2 (b-a)\sigma(S_a)^2, \tag{5.1}$$

$$\mathbb{E}^{\mathcal{F}_a}\left[\int_a^b |\sigma(S_t)-\sigma(S_a)|^2\,\mathrm{d}t\right] \le c_{(5.2)}^2 (b-a)^2\sigma(S_a)^2. \tag{5.2}$$

PROOF. See the Supplementary Material [37], subsection D.1. □

Proof of Proposition 3.6. For $\vartheta \in \mathcal{A}(S)$ and $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\mathrm{det}}$, we define the process $\langle \vartheta, \tau\rangle$, which is adapted, has continuous and nondecreasing paths on $[0, T]$, by

$$\langle \vartheta, \tau\rangle_t := \sum_{i=1}^n \int_{t_{i-1}\wedge t}^{t_i \wedge t} |\vartheta_u - \vartheta_{t_{i-1}}|^2 \sigma(S_u)^2\,\mathrm{d}u. \tag{5.3}$$

For $a \in [0, T)$, applying conditional Itô's isometry and Hölder's inequality yields, a.s.,

$$\begin{aligned}
&\mathbb{E}^{\mathcal{F}_a}\bigl[\bigl|E_T^{\mathrm{Rm}}(\vartheta,\tau) - E_a^{\mathrm{Rm}}(\vartheta,\tau)\bigr|^2\bigr] \\
&\le 3\mathbb{E}^{\mathcal{F}_a}\left[\int_a^T \left|\vartheta_{u-} - \sum_{i=1}^n \vartheta_{t_{i-1}-}\mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2 \sigma(S_{u-})^2\right. \\
&\qquad \left.\times \left(|a_u^Z|^2 + |j_u^Z|^2 + \int_a^T |b_r^Z|^2\,\mathrm{d}r\right)\mathrm{d}u\right] \\
&\le 3\bigl(|a_{(2.5)}^Z|^2 + |j_{(2.6)}^Z|^2 + |b_{(2.6)}^Z|^2\bigr) \\
&\qquad \times \mathbb{E}^{\mathcal{F}_a}\left[\int_a^T \left|\vartheta_{u-} - \sum_{i=1}^n \vartheta_{t_{i-1}-}\mathbb{1}_{(t_{i-1},t_i]}(u)\right|^2 \sigma(S_{u-})^2\,\mathrm{d}u\right] \\
&= 3\bigl(|a_{(2.5)}^Z|^2 + |j_{(2.6)}^Z|^2 + |b_{(2.6)}^Z|^2\bigr)\mathbb{E}^{\mathcal{F}_a}\bigl[\langle\vartheta,\tau\rangle_T - \langle\vartheta,\tau\rangle_a\bigr],
\end{aligned} \tag{5.4}$$

where the equality comes from the fact that the number of discontinuities of a càdlàg function is at most countable and $\vartheta \in \mathcal{A}(S)$ has no fixed-time discontinuity. Recall from Remark 2.2 that one can use deterministic times instead of stopping times in the definition of $\|\cdot\|_{\mathrm{bmo}_2^\Phi(\mathbb{P})}$. Therefore, Proposition 3.6 is a direct consequence of (5.4) and the following lemma.

LEMMA 5.2. *Let Assumption* 3.4 *hold for some* $\theta \in (0, 1]$. *Then there exists a constant* $c_{(5.5)} > 0$ *such that for any* $\tau \in \mathcal{T}_{\mathrm{det}}$ *and any* $a \in [0, T)$, *a.s.*,

$$\mathbb{E}^{\mathcal{F}_a}\bigl[\langle\vartheta,\tau\rangle_T - \langle\vartheta,\tau\rangle_a\bigr] \le c_{(5.5)}^2 \|\tau\|_\theta \Phi_a^2. \tag{5.5}$$

Consequently, $\|\langle\vartheta,\tau\rangle\|_{\mathrm{BMO}_1^{\Phi^2}(\mathbb{P})} \le c_{(5.5)}^2 \|\tau\|_\theta$.

PROOF. See the Supplementary Material [37], subsection D.3. □

5.2. *Proofs of results in Sections* 3.2 *to* 3.4. Let $\varepsilon > 0$, $\kappa \ge 0$ and recall $\rho(\varepsilon,\kappa) = (\rho_i(\varepsilon,\kappa))_{i\ge 0}$ in Definition 3.8. Due to (3.5) and the assumption $\sigma(S_-) > 0$, it holds that

$$|\Delta S| > \sigma(S_-)\varepsilon(T-\cdot)^\kappa \quad \Leftrightarrow \quad |\Delta Z| > \varepsilon(T-\cdot)^\kappa.$$

Hence, we derive from (3.6) the relations

$$\rho_i(\varepsilon,\kappa) = \inf\bigl\{T \ge t > \rho_{i-1}(\varepsilon,\kappa) : |\Delta Z_t| > \varepsilon(T-t)^\kappa\bigr\} \wedge T, \quad i \ge 1. \tag{5.6}$$

Since Z is càdlàg and the underlying filtration satisfies the usual conditions (right continuity and completeness), it implies that $\rho_i(\varepsilon,\kappa)$ are stopping times satisfying $\rho_{i-1}(\varepsilon,\kappa) < \rho_i(\varepsilon,\kappa)$ for $1 \le i \le \mathcal{N}_{(3.7)}(\varepsilon,\kappa)$.

For a nonnegative Borel function f defined on $\mathbb{R}$, we denote

$$\|f(z) \star \nu\|_{L_\infty(\mathbb{P}\otimes\lambda)} := \left\|(\omega,t) \mapsto \int_{\mathbb{R}} f(z)\nu_t(\omega,\mathrm{d}z)\right\|_{L_\infty(\Omega\times[0,T],\mathbb{P}\otimes\lambda)} \in [0,\infty].$$

Then, condition (3.14) is rewritten as

$$\|z^2 \star \nu\|_{L_\infty(\mathbb{P}\otimes\lambda)} < \infty. \tag{5.7}$$

LEMMA 5.3. *Let $\varepsilon > 0$, $\kappa \ge 0$ be real numbers. Then, for any $\alpha \in [0, \frac{1}{\kappa})$, one has*

$$\|\mathcal{N}_{(3.7)}(\varepsilon,\kappa)\|_{L_2(\mathbb{P})} \le 1 + \sqrt{c_{(5.8)}} + c_{(5.8)}, \tag{5.8}$$

where $c_{(5.8)} := T\|\mathbb{1}_{\{|z|>1\}} \star \nu\|_{L_\infty(\mathbb{P}\otimes\lambda)} + \varepsilon^{-\alpha}\sup_{r\in(0,1)}\|r^\alpha \mathbb{1}_{\{r<|z|\le 1\}} \star \nu\|_{L_\infty(\mathbb{P}\otimes\lambda)} \frac{T^{1-\alpha\kappa}}{1-\alpha\kappa}$.

PROOF. We may assume that $c_{(5.8)} < \infty$, otherwise (5.8) is trivial.

Step 1. We show that $A := \int_0^T\int_{\mathbb{R}} \mathbb{1}_{\{|z|>\varepsilon(T-t)^\kappa\}}\pi_Z(\mathrm{d}t,\mathrm{d}z) \le c_{(5.8)}$ a.s. Indeed, one has

$$A = \int_0^T\int_{\mathbb{R}} \mathbb{1}_{\{|z|>1\vee(\varepsilon(T-t)^\kappa)\}}\pi_Z(\mathrm{d}t,\mathrm{d}z) + \int_0^T\int_{\mathbb{R}} \mathbb{1}_{\{1\ge|z|>\varepsilon(T-t)^\kappa\}}\pi_Z(\mathrm{d}t,\mathrm{d}z),$$

where the first term on the right-hand side is upper bounded by $T\|\mathbb{1}_{\{|z|>1\}} \star \nu\|_{L_\infty(\mathbb{P}\otimes\lambda)}$ a.s. Let us denote

$$c_{(5.9)} := \sup_{r\in(0,1)} \|r^\alpha \mathbb{1}_{\{r<|z|\le 1\}} \star \nu\|_{L_\infty(\mathbb{P}\otimes\lambda)} < \infty. \tag{5.9}$$

By a standard approximation argument using a countable dense set of $(0,1)$, we infer that

$$\int_{r<|z|\le 1} \nu_t(\omega,\mathrm{d}z) \le c_{(5.9)} r^{-\alpha} \quad \forall r \in (0,1),$$

for $\mathbb{P}\otimes\lambda$-a.e. $(\omega,t) \in \Omega\times[0,T]$. For the second term in the decomposition of A, using Fubini's theorem we get that, a.s.,

$$\begin{aligned}
&\int_0^T\int_{\mathbb{R}} \mathbb{1}_{\{1\ge|z|>\varepsilon(T-t)^\kappa\}}\pi_Z(\mathrm{d}t,\mathrm{d}z) \\
&\quad = \int_{\{(t,z)\in[0,T]\times\mathbb{R}:\,\varepsilon(T-t)^\kappa<|z|\le 1\}} \nu_t(\mathrm{d}z)\,\mathrm{d}t \\
&\quad \le c_{(5.9)}\varepsilon^{-\alpha}\int_0^T (T-t)^{-\alpha\kappa}\,\mathrm{d}t = c_{(5.9)}\varepsilon^{-\alpha}\frac{T^{1-\alpha\kappa}}{1-\alpha\kappa}.
\end{aligned}$$

Step 2. Combining *Step 1* with [25], Chapter II, Proposition 1.28, allows us to write, a.s.,

$$\int_0^T\int_{\mathbb{R}} \mathbb{1}_{\{|z|>\varepsilon(T-t)^\kappa\}}N_Z(\mathrm{d}t,\mathrm{d}z) = \int_0^T\int_{\mathbb{R}} \mathbb{1}_{\{|z|>\varepsilon(T-t)^\kappa\}}[(N_Z-\pi_Z)(\mathrm{d}t,\mathrm{d}z) + \pi_Z(\mathrm{d}t,\mathrm{d}z)].$$

Since $\mathcal{N}_{(3.7)}(\varepsilon,\kappa) \le 1 + \int_0^T\int_{\mathbb{R}} \mathbb{1}_{\{|z|>\varepsilon(T-t)^\kappa\}}N_Z(\mathrm{d}t,\mathrm{d}z)$ by (5.6), we have

$$\begin{aligned}
\|\mathcal{N}_{(3.7)}(\varepsilon,\kappa)\|_{L_2(\mathbb{P})} &\le 1 + \left\|\int_0^T\int_{\mathbb{R}} \mathbb{1}_{\{|z|>\varepsilon(T-t)^\kappa\}}N_Z(\mathrm{d}t,\mathrm{d}z)\right\|_{L_2(\mathbb{P})} \\
&\le 1 + \left\|\int_0^T\int_{\mathbb{R}} \mathbb{1}_{\{|z|>\varepsilon(T-t)^\kappa\}}(N_Z-\pi_Z)(\mathrm{d}t,\mathrm{d}z)\right\|_{L_2(\mathbb{P})}
\end{aligned}$$

$$
\begin{aligned}
&+ \left\| \int_0^T \int_{\mathbb{R}} \mathbb{1}_{\{|z| > \varepsilon (T-t)^\kappa\}} \pi_Z(\mathrm{d}t, \mathrm{d}z) \right\|_{L_2(\mathbb{P})} \\
&= 1 + \left\| \int_0^T \int_{\mathbb{R}} \mathbb{1}_{\{|z| > \varepsilon (T-t)^\kappa\}} \pi_Z(\mathrm{d}t, \mathrm{d}z) \right\|_{L_1(\mathbb{P})}^{\frac{1}{2}} \\
&\quad + \left\| \int_0^T \int_{\mathbb{R}} \mathbb{1}_{\{|z| > \varepsilon (T-t)^\kappa\}} \pi_Z(\mathrm{d}t, \mathrm{d}z) \right\|_{L_2(\mathbb{P})} \\
&\le 1 + \sqrt{c_{(5.8)}} + c_{(5.8)},
\end{aligned}
$$

where one uses [25], Chapter II, Theorem 1.33(a), to derive the equality. □

LEMMA 5.4. *Let* $\alpha \in [0, \infty)$. *Assume that* μ *is a Borel measure on* $[-1, 1]$ *with* $\mu(\{0\}) = 0$. *If* $\sup_{r \in (0,1)} r^\alpha \int_{r < |x| \le 1} \mu(\mathrm{d}x) \le c_{\mu,\alpha} < \infty$, *then for* $\gamma > \alpha$ *one has*

$$
\int_{|x| \le r} |x|^\gamma \mu(\mathrm{d}x) \le \frac{c_{\mu,\alpha} 2^\gamma}{1 - 2^{\alpha - \gamma}} r^{\gamma - \alpha} \quad \textit{for any } r \in (0, 1],
$$

and for $0 < \gamma \le \alpha$ *one has*

$$
\int_{r < |x| \le 1} |x|^\gamma \mu(\mathrm{d}x) \le \begin{cases} c_{\mu,\alpha} 2^\alpha (1 - \log r) & \textit{if } \gamma = \alpha, \\ \dfrac{c_{\mu,\alpha} 2^{2\alpha - \gamma}}{2^{\alpha - \gamma} - 1} r^{\gamma - \alpha} & \textit{if } \gamma \in (0, \alpha) \end{cases} \quad \textit{for all } r \in (0, 1].
$$

PROOF. The proof is analogous to [17], Lemma 9.20, and is provided in the Supplementary Material [37], subsection D.4. □

5.2.1. *Proof of Theorem* 3.10. Recall $\kappa = \frac{1-\theta}{2} \in [0, \frac{1}{2})$.

Step 1. We handle the correction term in (3.8) and the corresponding error. For $\varepsilon > 0$,

$$
\begin{aligned}
\mathbb{E} \int_0^T \int_{\mathbb{R}} |z| \mathbb{1}_{\{|z| > \varepsilon (T-t)^\kappa\}} \nu_t(\mathrm{d}z) \, \mathrm{d}t &\le \varepsilon^{-1} \mathbb{E} \int_0^T (T-t)^{-\kappa} \int_{\mathbb{R}} z^2 \nu_t(\mathrm{d}z) \, \mathrm{d}t \\
&\le \varepsilon^{-1} \frac{T^{1-\kappa}}{1-\kappa} \| z^2 \star \nu \|_{L_\infty(\mathbb{P} \otimes \lambda)} < \infty,
\end{aligned}
$$

where the finiteness holds due to (5.7). This allows us to decompose

$$
\int_0^\cdot \int_{\mathbb{R}_0} z (N_Z - \pi_Z)(\mathrm{d}u, \mathrm{d}z) = Z^{\varepsilon,1} + Z^{\varepsilon,2} - \gamma^\varepsilon,
$$

where

$$
\begin{aligned}
Z^{\varepsilon,1} &:= \int_0^\cdot \int_{\mathbb{R}_0} z \mathbb{1}_{\{|z| \le \varepsilon (T-u)^\kappa\}} (N_Z - \pi_Z)(\mathrm{d}u, \mathrm{d}z), \\
Z^{\varepsilon,2} &:= \int_0^\cdot \int_{\mathbb{R}} z \mathbb{1}_{\{|z| > \varepsilon (T-u)^\kappa\}} N_Z(\mathrm{d}u, \mathrm{d}z), \\
\gamma^\varepsilon &:= \int_0^\cdot \int_{\mathbb{R}} z \mathbb{1}_{\{|z| > \varepsilon (T-u)^\kappa\}} \nu_u(\mathrm{d}z) \, \mathrm{d}u.
\end{aligned}
$$

Recall ϑ^τ in Definition 3.9. Since (5.7) holds in our context, applying Lemma 5.3 with $\alpha = 2$ yields $\mathcal{N}_{(3.7)}(\varepsilon, \kappa) < \infty$ a.s. Hence, outside a set of probability zero, we have that, for all $t \in [0, T]$,

$$
\sum_{\rho_i(\varepsilon,\kappa) \in [0,t] \cap [0,T)} \left(\vartheta_{\rho_i(\varepsilon,\kappa)-} - \vartheta^\tau_{\rho_i(\varepsilon,\kappa)} \right) \Delta S_{\rho_i(\varepsilon,\kappa)} = \int_{[0,t] \cap [0,T)} \left(\vartheta_{u-} - \vartheta^\tau_u \right) \sigma(S_{u-}) \, \mathrm{d}Z^{\varepsilon,2}_u.
$$

By the representation of Z in (2.3), one can decompose

$$\begin{aligned} \mathrm{d}S_t = \sigma(S_{t-})\,\mathrm{d}Z_t &= \sigma(S_{t-})\Big(\mathrm{d}Z^{\mathrm{c}}_t + b^Z_t\,\mathrm{d}t + \int_{\mathbb{R}_0} z(N_Z - \pi_Z)(\mathrm{d}t, \mathrm{d}z)\Big) \\ &= \sigma(S_{t-})\big(\mathrm{d}Z^{\mathrm{c}}_t + b^Z_t\,\mathrm{d}t + \mathrm{d}Z^{\varepsilon,1}_t + \mathrm{d}Z^{\varepsilon,2}_t - \mathrm{d}\gamma^{\varepsilon}_t\big). \end{aligned}$$

We get from the arguments above, together with the fact $\Delta Z^{\varepsilon,2}_T = \Delta Z_T = 0$ a.s., that

$$\begin{aligned} E^{\mathrm{adap}}(\vartheta,\tau|\varepsilon,\kappa) &= \int_0^{\cdot} (\vartheta_{u-} - \vartheta^{\tau}_u)\,\mathrm{d}S_u - \sum_{\rho_i(\varepsilon,\kappa)\in[0,\cdot]\cap[0,T)} \big(\vartheta_{\rho_i(\varepsilon,\kappa)-} - \vartheta^{\tau}_{\rho_i(\varepsilon,\kappa)}\big)\Delta S_{\rho_i(\varepsilon,\kappa)} \\ &= \int_0^{\cdot} (\vartheta_{u-} - \vartheta^{\tau}_u)\sigma(S_{u-})\big(\mathrm{d}Z^{\mathrm{c}}_u + b^Z_u\,\mathrm{d}u + \mathrm{d}Z^{\varepsilon,1}_u + \mathrm{d}Z^{\varepsilon,2}_u - \mathrm{d}\gamma^{\varepsilon}_u\big) \\ &\quad - \int_0^{\cdot} (\vartheta_{u-} - \vartheta^{\tau}_u)\sigma(S_{u-})\,\mathrm{d}Z^{\varepsilon,2}_u \\ &= E^{\mathrm{C}}(\vartheta,\tau|\varepsilon,\kappa) + E^{\mathrm{S}}(\vartheta,\tau|\varepsilon,\kappa) - E^{\mathrm{D}}(\vartheta,\tau|\varepsilon,\kappa), \end{aligned} \tag{5.10}$$

where the "continuous part", "small jump part" and "drift part" are given by

$$\begin{aligned} E^{\mathrm{C}}(\vartheta,\tau|\varepsilon,\kappa) &:= \int_0^{\cdot} (\vartheta_{u-} - \vartheta^{\tau}_u)\sigma(S_{u-})\big(\mathrm{d}Z^{\mathrm{c}}_u + b^Z_u\,\mathrm{d}u\big), \\ E^{\mathrm{S}}(\vartheta,\tau|\varepsilon,\kappa) &:= \int_0^{\cdot} (\vartheta_{u-} - \vartheta^{\tau}_u)\sigma(S_{u-})\,\mathrm{d}Z^{\varepsilon,1}_u, \\ E^{\mathrm{D}}(\vartheta,\tau|\varepsilon,\kappa) &:= \int_0^{\cdot} (\vartheta_{u-} - \vartheta^{\tau}_u)\sigma(S_{u-}) \int_{\mathbb{R}} z\mathbb{1}_{\{|z|>\varepsilon(T-u)^{\kappa}\}}\nu_u(\mathrm{d}z)\,\mathrm{d}u. \end{aligned}$$

The triangle inequality applied to (5.10) gives

$$\big\|E^{\mathrm{adap}}(\vartheta,\tau|\varepsilon,\kappa)\big\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \le \sum_{i\in\{\mathrm{S},\mathrm{C},\mathrm{D}\}} \big\|E^{i}(\vartheta,\tau|\varepsilon,\kappa)\big\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})}. \tag{5.11}$$

Step 2. We preliminary investigate the right-hand side of (5.11).

Step 2.1. Consider $E^{\mathrm{C}}(\vartheta,\tau|\varepsilon,\kappa)$. We apply the conditional Itô isometry for the martingale component and apply the Cauchy–Schwarz inequality for the finite variation component of $E^{\mathrm{C}}(\vartheta,\tau|\varepsilon,\kappa)$ to derive that, for $a\in[0,T)$, a.s.,

$$\begin{aligned} &\sqrt{\mathbb{E}^{\mathcal{F}_a}\big[\big|E^{\mathrm{C}}_T(\vartheta,\tau|\varepsilon,\kappa) - E^{\mathrm{C}}_a(\vartheta,\tau|\varepsilon,\kappa)\big|^2\big]} \\ &\quad\le \sqrt{\mathbb{E}^{\mathcal{F}_a}\Big[\int_a^T |\vartheta_{u-} - \vartheta^{\tau}_u|^2\sigma(S_{u-})^2\,\mathrm{d}\langle Z^{\mathrm{c}}\rangle_u\Big]} \\ &\qquad + \sqrt{\mathbb{E}^{\mathcal{F}_a}\Big[\int_a^T |\vartheta_{u-} - \vartheta^{\tau}_u|^2\sigma(S_{u-})^2\,\mathrm{d}u \int_a^T |b^Z_u|^2\,\mathrm{d}u\Big]} \\ &\quad\le \big(a^Z_{(2.5)} + b^Z_{(2.6)}\big)\sqrt{\mathbb{E}^{\mathcal{F}_a}\big[\langle\vartheta,\tau\rangle_T - \langle\vartheta,\tau\rangle_a\big]} \\ &\quad\le \big(a^Z_{(2.5)} + b^Z_{(2.6)}\big)c_{(5.5)}\sqrt{\|\tau\|_{\theta}}\,\Phi_a \\ &\quad\le \big(a^Z_{(2.5)} + b^Z_{(2.6)}\big)c_{(5.5)}\sqrt{\|\tau\|_{\theta}}\,\overline{\Phi}_a, \end{aligned}$$

where $\langle\vartheta,\tau\rangle$ is given in (5.3), and where we use the fact that a càdlàg function has at most countably many discontinuities for the second inequality. Since $E^{\mathrm{C}}(\vartheta,\tau|\varepsilon,\kappa)$ has continuous paths, it implies that

$$\begin{aligned} \big\|E^{\mathrm{C}}(\vartheta,\tau|\varepsilon,\kappa)\big\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} &= \big\|E^{\mathrm{C}}(\vartheta,\tau|\varepsilon,\kappa)\big\|_{\mathrm{bmo}_2^{\overline{\Phi}}(\mathbb{P})} \\ &\le c_{(5.5)}\big(a^Z_{(2.5)} + b^Z_{(2.6)}\big)\sqrt{\|\tau\|_{\theta}}. \end{aligned} \tag{5.12}$$

Step 2.2. Consider $E^{\mathrm{S}}(\vartheta, \tau|\varepsilon, \kappa)$. Since $\overline{\Phi} \in \mathcal{SM}_2(\mathbb{P})$ by assumption, Lemma 2.3 shows

$$\|E^{\mathrm{S}}(\vartheta, \tau|\varepsilon, \kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \sim_{c_{(5.13)}} \|E^{\mathrm{S}}(\vartheta, \tau|\varepsilon, \kappa)\|_{\mathrm{bmo}_2^{\overline{\Phi}}(\mathbb{P})} + |\Delta E^{\mathrm{S}}(\vartheta, \tau|\varepsilon, \kappa)|_{\overline{\Phi}}. \tag{5.13}$$

Recall that Θ is nondecreasing by assumption. Since ϑ, $\sigma(S)$ and Φ are càdlàg on $[0, T)$, one can find an Ω_0 with $\mathbb{P}(\Omega_0) = 1$ such that for $\omega \in \Omega_0$ we have

$$|\vartheta_t - \vartheta_s|\sigma(S_t) \le 2c_{(3.2)}(T-t)^{-\kappa}\Phi_t \quad \forall 0 \le s < t < T.$$

Due to (2.4), one has $\pi_Z(\omega, \{t\} \times \mathbb{R}_0) = 0$ for any $(\omega, t) \in \Omega \times [0, T]$. Then it holds that

$$|\Delta Z_t^{\varepsilon,1}| = \left|\int_{\mathbb{R}_0} z\mathbb{1}_{\{|z|\le\varepsilon(T-t)^\kappa\}}N_Z(\{t\}, \mathrm{d}z) - \int_{\mathbb{R}_0} z\mathbb{1}_{\{|z|\le\varepsilon(T-t)^\kappa\}}\pi_Z(\{t\}, \mathrm{d}z)\right| \le \varepsilon(T-t)^\kappa$$

for all $t \in [0, T]$ a.s. Moreover, since $\Delta E^{\mathrm{S}}(\vartheta, \tau|\varepsilon, \kappa) = (\vartheta_- - \vartheta^\tau)\sigma(S_-)\Delta Z^{\varepsilon,1}$, there is an Ω_1 with $\mathbb{P}(\Omega_1) = 1$ (with keeping $\Delta Z_T^{\varepsilon,1} = 0$ a.s. in mind) such that for all $(\omega, t) \in \Omega_1 \times [0, T]$,

$$\begin{aligned}|\Delta E_t^{\mathrm{S}}(\vartheta, \tau|\varepsilon, \kappa)| &= |(\vartheta_{t-} - \vartheta_t^\tau)\sigma(S_{t-})\Delta Z_t^{\varepsilon,1}| \le 2c_{(3.2)}(T-t)^{-\kappa}\Phi_{t-}\varepsilon(T-t)^\kappa \\ &= 2c_{(3.2)}\varepsilon\Phi_{t-} \le 2c_{(3.2)}\varepsilon\overline{\Phi}_t.\end{aligned}$$

According to the definition of $|\cdot|_{\overline{\Phi}}$ given in Lemma 2.3, one then gets

$$|\Delta E^{\mathrm{S}}(\vartheta, \tau|\varepsilon, \kappa)|_{\overline{\Phi}} \le 2c_{(3.2)}\varepsilon. \tag{5.14}$$

Let us continue with $\|E^{\mathrm{S}}(\vartheta, \tau|\varepsilon, \kappa)\|_{\mathrm{bmo}_2^{\overline{\Phi}}(\mathbb{P})}$. For $a \in [0, T)$, we have, a.s.,

$$\begin{aligned}&\mathbb{E}^{\mathcal{F}_a}\big[|E_T^{\mathrm{S}}(\vartheta, \tau|\varepsilon, \kappa) - E_a^{\mathrm{S}}(\vartheta, \tau|\varepsilon, \kappa)|^2\big] \\ &\quad = \mathbb{E}^{\mathcal{F}_a}\left[\int_a^T |\vartheta_{u-} - \vartheta_u^\tau|^2\sigma(S_{u-})^2\int_{\mathbb{R}}\mathbb{1}_{\{|z|\le\varepsilon(T-u)^\kappa\}}z^2\nu_u(\mathrm{d}z)\,\mathrm{d}u\right] \\ &\quad \le \|\mathbb{1}_{\{|z|\le\varepsilon T^\kappa\}}z^2 \star \nu\|_{L_\infty(\mathbb{P}\otimes\lambda)}\mathbb{E}^{\mathcal{F}_a}\left[\int_a^T |\vartheta_{u-} - \vartheta_u^\tau|^2\sigma(S_{u-})^2\,\mathrm{d}u\right] \\ &\quad = \|\mathbb{1}_{\{|z|\le\varepsilon T^\kappa\}}z^2 \star \nu\|_{L_\infty(\mathbb{P}\otimes\lambda)}\mathbb{E}^{\mathcal{F}_a}[\langle\vartheta, \tau\rangle_T - \langle\vartheta, \tau\rangle_a] \\ &\quad \le \|\mathbb{1}_{\{|z|\le\varepsilon T^\kappa\}}z^2 \star \nu\|_{L_\infty(\mathbb{P}\otimes\lambda)}c_{(5.5)}^2\|\tau\|_\theta\overline{\Phi}_a^2.\end{aligned} \tag{5.15}$$

Combing (5.14) and (5.15) with (5.13) we obtain for $c_{(5.16)} := c_{(5.13)}((2c_{(3.2)}) \vee c_{(5.5)})$ that

$$\|E^{\mathrm{S}}(\vartheta, \tau|\varepsilon, \kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \le c_{(5.16)}\Big(\varepsilon + \sqrt{\|\mathbb{1}_{\{|z|\le\varepsilon T^\kappa\}}z^2 \star \nu\|_{L_\infty(\mathbb{P}\otimes\lambda)}}\sqrt{\|\tau\|_\theta}\Big). \tag{5.16}$$

Step 2.3. Consider $E^{\mathrm{D}}(\vartheta, \tau|\varepsilon, \kappa)$. Since $E^{\mathrm{D}}(\vartheta, \tau|\varepsilon, \kappa)$ is continuous, it holds that

$$\|E^{\mathrm{D}}(\vartheta, \tau|\varepsilon, \kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} = \|E^{\mathrm{D}}(\vartheta, \tau|\varepsilon, \kappa)\|_{\mathrm{bmo}_2^{\overline{\Phi}}(\mathbb{P})}.$$

Then, for any $a \in [0, T)$, we use the Cauchy–Schwarz inequality to get, a.s.,

$$\begin{aligned}&\mathbb{E}^{\mathcal{F}_a}\big[|E_T^{\mathrm{D}}(\vartheta, \tau|\varepsilon, \kappa) - E_a^{\mathrm{D}}(\vartheta, \tau|\varepsilon, \kappa)|^2\big] \\ &\quad \le \mathbb{E}^{\mathcal{F}_a}\left[\left(\int_a^T\left|\int_{\mathbb{R}} z\mathbb{1}_{\{|z|>\varepsilon(T-u)^\kappa\}}\nu_u(\mathrm{d}z)\right|^2\mathrm{d}u\right)\left(\int_a^T |\vartheta_{u-} - \vartheta_u^\tau|^2\sigma(S_{u-})^2\,\mathrm{d}u\right)\right] \\ &\quad =: \mathbb{E}^{\mathcal{F}_a}[\mathrm{I}_{(5.17)}\mathrm{II}_{(5.17)}].\end{aligned} \tag{5.17}$$

(1) We now exploit the condition (3.10). By a standard approximation argument using a countable dense set of $(0, 1)$, we get that, for $\mathbb{P} \otimes \lambda$-a.e. $(\omega, u) \in \Omega \times [0, T]$,

$$\int_{r<|z|\le1}\nu_u(\omega, \mathrm{d}z) \le c_{(3.10)}r^{-\alpha} \quad \forall r \in (0, 1). \tag{5.18}$$

Let us first examine the right-hand side of (5.16). For any $(\omega, u) \in \Omega \times [0, T]$ such that (5.18) is satisfied and that $\int_{\mathbb{R}} z^2 \nu_u(\omega, \mathrm{d}z) \le \|z^2 \star \nu\|_{L_\infty(\mathbb{P}\otimes\lambda)} < \infty$, there exists a $c_{(5.19)} > 0$ independent of ω, u, ε such that

$$\int_{|z|\le\varepsilon T^\kappa} z^2 \nu_u(\omega, \mathrm{d}z) \le c_{(5.19)}\big((\varepsilon T^\kappa) \wedge 1\big)^{2-\alpha}. \tag{5.19}$$

Indeed, if $\alpha < 2$ then we apply Lemma 5.4 with $\mu(\cdot) = \nu_u(\omega, \cdot)$ and $\gamma = 2 > \alpha$ to get (5.19). If $\alpha = 2$ or if $\varepsilon T^\kappa > 1$, then (5.19) obviously holds. Thus,

$$\|\mathbb{1}_{\{|z|\le\varepsilon T^\kappa\}} z^2 \star \nu\|_{L_\infty(\mathbb{P}\otimes\lambda)} \le c_{(5.19)}\big((\varepsilon T^\kappa) \wedge 1\big)^{2-\alpha} \le c_{(5.19)}\big(1 + T^\kappa\big)^{2-\alpha}(\varepsilon \wedge 1)^{2-\alpha}.$$

Then it follows from (5.16) that

$$\|E^{\mathrm{S}}(\vartheta, \tau|\varepsilon, \kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \le c_{(5.20)}\big(\varepsilon + (\varepsilon \wedge 1)^{1-\frac{\alpha}{2}}\sqrt{\|\tau\|_\theta}\big) \tag{5.20}$$

for some constant $c_{(5.20)} > 0$ independent of ε and τ.

We continue with $\mathrm{I}_{(5.17)}$ and $\mathrm{II}_{(5.17)}$. Let $(\omega, u) \in \Omega \times [0, T]$ such that $0 < \varepsilon(T-u)^\kappa \le 1$ and (5.18) holds. We first apply Lemma 5.4 and then use Fubini's theorem to get that, a.s.,

$$\begin{aligned}
&\sqrt{\int_0^T \left|\int_{\varepsilon(T-u)^\kappa<|z|\le 1} |z|\nu_u(\mathrm{d}z)\right|^2 \mathrm{d}u} \\
&\le 2c_{(3.10)} \begin{cases} \dfrac{\sqrt{T}}{1-2^{\alpha-1}} & \text{if } \alpha \in (0,1), \\ \sqrt{T}\Big(\log^+\Big(\dfrac{1}{\varepsilon}\Big) + 1\Big) + \kappa\sqrt{\displaystyle\int_0^T \log^2(T-u)\,\mathrm{d}u} & \text{if } \alpha = 1, \\ \Big[\dfrac{2^{2\alpha-2}}{2^{\alpha-1}-1}\sqrt{\displaystyle\int_0^T (T-u)^{2\kappa(1-\alpha)}\,\mathrm{d}u}\Big]\varepsilon^{1-\alpha} & \text{if } \alpha \in (1,2] \end{cases} \\
&\le 2c_{(3.10)}c_{\alpha,\kappa,T} \begin{cases} 1 & \text{if } \alpha \in (0,1), \\ \log^+\Big(\dfrac{1}{\varepsilon}\Big) + 1 & \text{if } \alpha = 1, \\ \varepsilon^{1-\alpha} & \text{if } \alpha \in (1,2] \end{cases}
\end{aligned}$$

for some constant $c_{\alpha,\kappa,T} > 0$ depending at most on α, κ, T, and where one notices that $2\kappa(1-\alpha) + 1 > 0$. For the first factor $\mathrm{I}_{(5.17)}$, the triangle inequality gives, a.s.,

$$\begin{aligned}
\sqrt{\mathrm{I}_{(5.17)}} &\le \sqrt{\int_a^T \left|\int_{|z|>1\vee(\varepsilon(T-u)^\kappa)} z\nu_u(\mathrm{d}z)\right|^2 \mathrm{d}u} + \sqrt{\int_a^T \left|\int_{\varepsilon(T-u)^\kappa<|z|\le 1} z\nu_u(\mathrm{d}z)\right|^2 \mathrm{d}u} \\
&\le \sqrt{T}\|\mathbb{1}_{\{|z|>1\}}|z| \star \nu\|_{L_\infty(\mathbb{P}\otimes\lambda)} + \sqrt{\int_0^T \left|\int_{\varepsilon(T-u)^\kappa<|z|\le 1} z\nu_u(\mathrm{d}z)\right|^2 \mathrm{d}u} \\
&\le c_{(5.21)}\big(1 + h(\varepsilon)\big)
\end{aligned} \tag{5.21}$$

for some constant $c_{(5.21)} = c_{(5.21)}(\alpha, \kappa, T, \nu) > 0$ and for

$$h(\varepsilon) = 1 \quad \text{if } \alpha \in (0,1), \qquad h(\varepsilon) = \log^+\Big(\frac{1}{\varepsilon}\Big) \quad \text{if } \alpha = 1, \qquad h(\varepsilon) = \varepsilon^{1-\alpha} \quad \text{if } \alpha \in (1,2].$$

For the second factor $\mathrm{II}_{(5.17)}$, we apply Lemma 5.2 to obtain, a.s.,

$$\mathbb{E}^{\mathcal{F}_a}[\mathrm{II}_{(5.17)}] = \mathbb{E}^{\mathcal{F}_a}\big[\langle\vartheta, \tau\rangle_T - \langle\vartheta, \tau\rangle_a\big] \le c_{(5.5)}^2\|\tau\|_\theta \Phi_a^2 \le c_{(5.5)}^2\|\tau\|_\theta \overline{\Phi}_a^2.$$

Hence,

$$\|E^{\mathrm{D}}(\vartheta,\tau|\varepsilon,\kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \le c_{(5.5)}c_{(5.21)}(1+h(\varepsilon))\sqrt{\|\tau\|_\theta}. \tag{5.22}$$

Eventually, we plug (5.12), (5.20) and (5.22) into (5.11) to derive (3.11).

(2) If (3.12) holds, then $\mathrm{I}_{(5.17)} \le 2T(\|\mathbb{1}_{\{|z|>1\}}|z|\star\nu\|^2_{L_\infty(\mathbb{P}\otimes\lambda)} + c^2_{(3.12)}) =: c^2_{(5.23)}$. Hence,

$$\|E^{\mathrm{D}}(\vartheta,\tau|\varepsilon,\kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \le c_{(5.23)}c_{(5.5)}\sqrt{\|\tau\|_\theta}. \tag{5.23}$$

Combining (5.12), (5.20) and (5.23) with (5.11) yields (3.13). □

5.2.2. *Proof of Theorem* 3.14. Recall $\kappa = \frac{1-\theta}{2} \in [0,\frac{1}{2})$. Let $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\det}$ and $\varepsilon > 0$. For $u \in [0,T)$, we define

$$D_{\varepsilon,\kappa}(u) := \int_{|z|>\varepsilon(T-u)^\kappa} z\nu_u(\mathrm{d}z).$$

The conclusion for $E^{\mathrm{adap}}(\vartheta,\tau|\varepsilon,\kappa)$ is shown by using again (5.11), where the estimate for the "small jump part" $E^{\mathrm{S}}(\vartheta,\tau|\varepsilon,\kappa)$ is taken from (5.20). Here, we focus on improving the estimate for the "drift part" $E^{\mathrm{D}}(\vartheta,\tau|\varepsilon,\kappa)$.

Step 1. We show that there is a constant $c_{(5.24)} > 0$ independent of τ and ε such that

$$\begin{aligned} D_{(5.24)} &:= \sup_{i=1,\dots,n}\ \sup_{r\in(t_{i-1},t_i]\cap[0,T)}\left[\frac{1}{(T-r)^\kappa}\int_r^{t_i}|D_{\varepsilon,\kappa}(u)|\,\mathrm{d}u\right] \\ &\le c_{(5.24)}\begin{cases} \|\tau\|_\theta & \text{if (3.17) holds,} \\ \left[1+\log^+\left(\frac{1}{\varepsilon}\right)+\log^+\left(\frac{1}{\|\tau\|_\theta}\right)\right]\|\tau\|_\theta & \text{if (3.16) holds with } \alpha=1, \\ \|\tau\|_\theta + \varepsilon^{1-\alpha}\|\tau\|_\theta^{1-\kappa(\alpha-1)} & \text{if (3.16) holds with } \alpha\in(1,2]. \end{cases} \end{aligned} \tag{5.24}$$

Indeed, since $\|z^2\star\nu\|_{L_\infty([0,T],\lambda)} = \|u\mapsto\int_{\mathbb{R}} z^2\nu_u(\mathrm{d}z)\|_{L_\infty([0,T],\lambda)} < \infty$ by assumption, we first get

$$|D_{\varepsilon,\kappa}(u)| \le \int_{|z|>1} z^2\nu_u(\mathrm{d}z) + \int_{1\ge|z|>\varepsilon(T-u)^\kappa}|z|\nu_u(\mathrm{d}z)$$

and then apply Lemma 5.4 to obtain a $c_{(5.25)} > 0$ not depending on ε such that, for λ-a.e. $u \in [0,T)$,

$$|D_{\varepsilon,\kappa}(u)| \le c_{(5.25)}\begin{cases} 1 & \text{if (3.17) holds,} \\ 1+\log^+\left(\frac{1}{\varepsilon}\right)+\log^+\left(\frac{1}{T-u}\right) & \text{if (3.16) holds with } \alpha=1, \\ 1+\varepsilon^{1-\alpha}(T-u)^{\kappa(1-\alpha)} & \text{if (3.16) holds with } \alpha\in(1,2]. \end{cases} \tag{5.25}$$

Case 1. If (3.17) holds, then (5.25) immediately implies

$$D_{(5.24)} \le c_{(5.25)}\sup_{i=1,\dots,n}\ \sup_{r\in(t_{i-1},t_i]\cap[0,T)}\left[\frac{t_i-r}{(T-r)^{2\kappa}}(T-r)^\kappa\right] \le c_{(5.25)}T^\kappa\|\tau\|_\theta.$$

Case 2. (3.16) holds with $\alpha = 1$. For $r \in [t_{i-1},t_i)$, $i = 1,\dots,n$, one has

$$\int_r^{t_i}\log^+\left(\frac{1}{T-u}\right)\mathrm{d}u \le (t_i-r)\log^+\left(\frac{1}{t_i-r}\right) + (t_i-r)\log\mathrm{e},$$

where we first integrate by parts the left-hand side and then use the inequality

$$b\log^+\left(\frac{1}{b}\right) - a\log^+\left(\frac{1}{a}\right) \le (b-a)\log^+\left(\frac{1}{b-a}\right), \quad 0<a<b.$$

Since $x \mapsto x\log^+(\frac{1}{x})$ is nondecreasing on $(0, \frac{1}{\mathrm{e}}]$ and $0 < \frac{1}{\mathrm{e}T^\theta}\frac{t_i - r}{(T-r)^{1-\theta}} \le \frac{\|\tau\|_\theta}{\mathrm{e}T^\theta} \le \frac{1}{\mathrm{e}}$ for any $r \in [t_{i-1}, t_i)$, we get

$$\begin{aligned}
&\frac{1}{(T-r)^\kappa}\int_r^{t_i}\log^+\Big(\frac{1}{T-u}\Big)\,\mathrm{d}u \\
&\quad\le \frac{1}{\mathrm{e}T^\theta}\frac{t_i - r}{(T-r)^{2\kappa}}\Big[\log^+\Big(\frac{\mathrm{e}T^\theta(T-r)^{2\kappa}}{t_i - r}\Big) + \log^+\Big(\frac{1}{\mathrm{e}T^\theta(T-r)^{2\kappa}}\Big) + \log \mathrm{e}\Big]\mathrm{e}T^\theta(T-r)^\kappa \\
&\quad\le T^\kappa\|\tau\|_\theta\log^+\Big(\frac{\mathrm{e}T^\theta}{\|\tau\|_\theta}\Big) + \|\tau\|_\theta\sup_{r\in[0,T)}\Big[(T-r)^\kappa\log^+\Big(\frac{1}{\mathrm{e}T^\theta(T-r)^{2\kappa}}\Big)\Big] + \|\tau\|_\theta T^\kappa\log \mathrm{e} \\
&\quad\le c_{T,\theta}\|\tau\|_\theta\Big(1 + \log^+\Big(\frac{1}{\|\tau\|_\theta}\Big)\Big)
\end{aligned}$$

for some constant $c_{T,\theta} > 0$ depending at most on T, θ. Hence,

$$D_{(5.24)} \le c_{(5.25)}\Big[T^\kappa\Big(1 + \log^+\Big(\frac{1}{\varepsilon}\Big)\Big) + c_{T,\theta}\Big(1 + \log^+\Big(\frac{1}{\|\tau\|_\theta}\Big)\Big)\Big]\|\tau\|_\theta.$$

Case 3. (3.16) holds with $\alpha \in (1, 2]$. Again, using (5.25) and keeping $\frac{1}{2} < 1 - \kappa(\alpha - 1) \le 1$ in mind we get that, for any $r \in (t_{i-1}, t_i] \cap [0, T)$ and any $i = 1, \dots, n$,

$$\begin{aligned}
&\frac{1}{(T-r)^\kappa}\int_r^{t_i}|D_{\varepsilon,\kappa}(u)|\,\mathrm{d}u \\
&\quad\le \frac{c_{(5.25)}}{(T-r)^\kappa}\Big[(t_i - r) + \varepsilon^{1-\alpha}\int_r^{t_i}(T-u)^{\kappa(1-\alpha)}\,\mathrm{d}u\Big] \\
&\quad= c_{(5.25)}\Big[\frac{t_i - r}{(T-r)^\kappa} + \frac{\varepsilon^{1-\alpha}}{1-\kappa(\alpha-1)}\frac{(T-r)^{1-\kappa(\alpha-1)} - (T-t_i)^{1-\kappa(\alpha-1)}}{(T-r)^\kappa}\Big] \\
&\quad\le c_{(5.25)}\Big[\frac{t_i - r}{(T-r)^\kappa} + \frac{\varepsilon^{1-\alpha}}{1-\kappa(\alpha-1)}\frac{(t_i - r)^{1-\kappa(\alpha-1)}}{(T-r)^\kappa}\Big] \\
&\quad= c_{(5.25)}\Big[\frac{t_i - r}{(T-r)^{2\kappa}}(T-r)^\kappa + \frac{\varepsilon^{1-\alpha}}{1-\kappa(\alpha-1)}\Big[\frac{t_i - r}{(T-r)^{2\kappa}}\Big]^{1-\kappa(\alpha-1)} \\
&\qquad\times (T-r)^{2\kappa(1-\kappa(\alpha-1))-\kappa}\Big] \\
&\quad\le c_{(5.25)}\Big[\|\tau\|_\theta T^\kappa + \frac{\varepsilon^{1-\alpha}}{1-\kappa(\alpha-1)}\|\tau\|_\theta^{1-\kappa(\alpha-1)}T^{2\kappa(1-\kappa(\alpha-1))-\kappa}\Big],
\end{aligned}$$

where one uses $2\kappa(1-\kappa(\alpha-1)) - \kappa \ge 0$ for the last inequality. Thus, the assertion follows.

Step 2. We examine the "drift part" $E^{\mathrm{D}}(\vartheta, \tau|\varepsilon, \kappa)$. We let $a \in [t_{k-1}, t_k)$, $k \in [1, n]$, and set $s_i := a \vee t_i$, $i = k-1, \dots, n$. Denote

$$D_{(5.26)} := \sup_{i=k,\dots,n}\ \sup_{r\in(s_{i-1},s_i]\cap[0,T)}\Big[\frac{1}{(T-r)^\kappa}\int_r^{s_i}|D_{\varepsilon,\kappa}(u)|\,\mathrm{d}u\Big]. \tag{5.26}$$

Recall from Assumption 3.13 that $\vartheta S = M + V$ where $V_t := \int_0^t v_u \,\mathrm{d}u$. Then one has, a.s.,

$$
\begin{aligned}
&\frac{1}{4}\mathbb{E}^{\mathcal{F}_a}\big[\big|E_T^{\mathrm{D}}(\vartheta,\tau|\varepsilon,\kappa) - E_a^{\mathrm{D}}(\vartheta,\tau|\varepsilon,\kappa)\big|^2\big] \\
&\quad = \frac{1}{4}\mathbb{E}^{\mathcal{F}_a}\Bigg[\Bigg|(\vartheta_a - \vartheta_{t_{k-1}})\int_a^{t_k} S_u D_{\varepsilon,\kappa}(u)\,\mathrm{d}u \\
&\qquad + \sum_{i=k}^{n}\int_{s_{i-1}}^{s_i}(\vartheta_u - \vartheta_{s_{i-1}})S_u D_{\varepsilon,\kappa}(u)\,\mathrm{d}u\Bigg|^2\Bigg] \\
&\quad \le \mathbb{E}^{\mathcal{F}_a}\Bigg[\Bigg|(\vartheta_a - \vartheta_{t_{k-1}})\int_a^{t_k} S_u D_{\varepsilon,\kappa}(u)\,\mathrm{d}u\Bigg|^2\Bigg] \\
&\qquad + \mathbb{E}^{\mathcal{F}_a}\Bigg[\Bigg|\sum_{i=k}^{n}\int_{s_{i-1}}^{s_i}(M_u - M_{s_{i-1}})D_{\varepsilon,\kappa}(u)\,\mathrm{d}u\Bigg|^2\Bigg] \\
&\qquad + \mathbb{E}^{\mathcal{F}_a}\Bigg[\Bigg|\sum_{i=k}^{n}\int_{s_{i-1}}^{s_i}(V_u - V_{s_{i-1}})D_{\varepsilon,\kappa}(u)\,\mathrm{d}u\Bigg|^2\Bigg] \\
&\qquad + \mathbb{E}^{\mathcal{F}_a}\Bigg[\Bigg|\sum_{i=k}^{n}\vartheta_{s_{i-1}}\int_{s_{i-1}}^{s_i}(S_u - S_{s_{i-1}})D_{\varepsilon,\kappa}(u)\,\mathrm{d}u\Bigg|^2\Bigg] \\
&\quad =: \mathrm{I}_{(5.27)} + \mathrm{II}_{(5.27)} + \mathrm{III}_{(5.27)} + \mathrm{IV}_{(5.27)}.
\end{aligned}
\tag{5.27}
$$

For $\mathrm{I}_{(5.27)}$, we make use of the growth property of ϑ and the monotonicity of Θ to get, a.s.,

$$
\begin{aligned}
\mathrm{I}_{(5.27)} &\le 4c_{(3.2)}^2(T-a)^{\theta-1}\Theta_a^2\Big[\int_a^{t_k}|D_{\varepsilon,\kappa}(u)|\,\mathrm{d}u\Big]^2\mathbb{E}^{\mathcal{F}_a}\Big[\sup_{u\in[a,t_k]} S_u^2\Big] \\
&\le 4D_{(5.26)}^2 c_{(3.2)}^2\mathbb{E}^{\mathcal{F}_a}\Big[\sup_{u\in[a,t_k]}\Theta_u^2 S_u^2\Big] \le 4D_{(5.26)}^2 c_{(3.2)}^2\mathbb{E}^{\mathcal{F}_a}\Big[\sup_{u\in[a,t_k]}\overline{\Phi}_u^2\Big] \\
&\le 4D_{(5.26)}^2 c_{(3.2)}^2\|\overline{\Phi}\|_{\mathcal{SM}_2(\mathbb{P})}^2\overline{\Phi}_a^2.
\end{aligned}
$$

For $\mathrm{II}_{(5.27)}$, using the orthogonality of martingale increments we find that the mixed terms in the square expansion vanish under the conditional expectation. Then, applying the stochastic Fubini theorem and the conditional Itô isometry we obtain, a.s.,

$$
\begin{aligned}
\mathrm{II}_{(5.27)} &= \sum_{i=k}^{n}\mathbb{E}^{\mathcal{F}_a}\Bigg[\Bigg|\int_{(s_{i-1},s_i]\cap(0,T)}(M_u - M_{s_{i-1}})D_{\varepsilon,\kappa}(u)\,\mathrm{d}u\Bigg|^2\Bigg] \\
&= \sum_{i=k}^{n}\mathbb{E}^{\mathcal{F}_a}\Bigg[\Bigg|\int_{(s_{i-1},s_i]\cap(0,T)}\Big(\int_{[r,s_i]}D_{\varepsilon,\kappa}(u)\,\mathrm{d}u\Big)\mathrm{d}M_r\Bigg|^2\Bigg] \\
&= \sum_{i=k}^{n}\mathbb{E}^{\mathcal{F}_a}\Bigg[\int_{(s_{i-1},s_i]\cap(0,T)}\Bigg|\int_{[r,s_i]}D_{\varepsilon,\kappa}(u)\,\mathrm{d}u\Bigg|^2\mathrm{d}\langle M\rangle_r\Bigg] \\
&\le \sup_{i=k,\dots,n}\ \sup_{r\in(s_{i-1},s_i]\cap[0,T)}\Bigg[\frac{1}{(T-r)^{1-\theta}}\Bigg|\int_{[r,s_i]}D_{\varepsilon,\kappa}(u)\,\mathrm{d}u\Bigg|^2\Bigg] \\
&\quad \times \mathbb{E}^{\mathcal{F}_a}\Bigg[\int_{(a,T)}(T-u)^{1-\theta}\,\mathrm{d}\langle M\rangle_u\Bigg] \\
&\le D_{(5.26)}^2\mathbb{E}^{\mathcal{F}_a}\Bigg[\int_{(a,T)}(T-u)^{1-\theta}\Upsilon(\cdot,\mathrm{d}u)\Bigg]
\end{aligned}
$$

$$\leq D^2_{(5.26)} c^2_{(3.3)} \Phi^2_a.$$

For $\mathrm{III}_{(5.27)}$, we use Fubini's theorem and Hölder's inequality to obtain, a.s.,

$$\begin{aligned}
\mathrm{III}_{(5.27)} &\leq \mathbb{E}^{\mathcal{F}_a}\left[\left|\sum_{i=k}^n \int_{(s_{i-1},s_i]\cap(0,T)} \left(\int_{s_{i-1}}^u |v_r|\,\mathrm{d}r\right)|D_{\varepsilon,\kappa}(u)|\,\mathrm{d}u\right|^2\right] \\
&\leq D^2_{(5.26)}\mathbb{E}^{\mathcal{F}_a}\left[\left|\int_{(a,T)} (T-r)^{\frac{1-\theta}{2}}|v_r|\,\mathrm{d}r\right|^2\right] \\
&\leq D^2_{(5.26)}(T-a)\mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,T)} (T-r)^{1-\theta} v_r^2\,\mathrm{d}r\right] \\
&\leq D^2_{(5.26)} T \mathbb{E}^{\mathcal{F}_a}\left[\int_{(a,T)} (T-r)^{1-\theta}\Upsilon(\cdot,\mathrm{d}r)\right] \\
&\leq D^2_{(5.26)} T c^2_{(3.3)} \Phi^2_a.
\end{aligned}$$

For $\mathrm{IV}_{(5.27)}$, we also exploit the martingale property of S, the monotonicity of Θ, and follow the same argument as for $\mathrm{II}_{(5.27)}$ to get, a.s.,

$$\begin{aligned}
\mathrm{IV}_{(5.27)} &= \sum_{i=k}^n \mathbb{E}^{\mathcal{F}_a}\left[\left|\vartheta_{s_{i-1}}\int_{s_{i-1}}^{s_i}(S_u - S_{s_{i-1}})D_{\varepsilon,\kappa}(u)\,\mathrm{d}u\right|^2\right] \\
&= \sum_{i=k}^n \mathbb{E}^{\mathcal{F}_a}\left[\vartheta^2_{s_{i-1}}\int_{(s_{i-1},s_i]}\left|\int_{[r,s_i]} D_{\varepsilon,\kappa}(u)\,\mathrm{d}u\right|^2 \mathrm{d}\langle S\rangle_r\right] \\
&\leq D^2_{(5.26)}\sum_{i=k}^n \mathbb{E}^{\mathcal{F}_a}\left[\vartheta^2_{s_{i-1}}\int_{(s_{i-1},s_i]}(T-r)^{1-\theta}\,\mathrm{d}\langle S\rangle_r\right] \\
&\leq D^2_{(5.26)}\|z^2\star\nu\|_{L_\infty([0,T],\lambda)}\sum_{i=k}^n \mathbb{E}^{\mathcal{F}_a}\left[(T-s_{i-1})^{1-\theta}\vartheta^2_{s_{i-1}}\mathbb{E}^{\mathcal{F}_{s_{i-1}}}\left[\int_{s_{i-1}}^{s_i} S_r^2\,\mathrm{d}r\right]\right] \\
&\leq D^2_{(5.26)}\|z^2\star\nu\|_{L_\infty([0,T],\lambda)} c^2_{(3.2)}\sum_{i=k}^n \mathbb{E}^{\mathcal{F}_a}\left[\int_{s_{i-1}}^{s_i}\Theta_r^2 S_r^2\,\mathrm{d}r\right] \\
&\leq D^2_{(5.26)}\|z^2\star\nu\|_{L_\infty([0,T],\lambda)} c^2_{(3.2)}(T-a)\|\overline{\Phi}\|^2_{\mathcal{SM}_2(\mathbb{P})}\overline{\Phi}^2_a,
\end{aligned}$$

where in order to obtain the second inequality we employ the assumption that $\mathrm{d}S_t = S_{t-}\,\mathrm{d}Z_t$ and $\mathrm{d}\langle Z\rangle_t = \int_{\mathbb{R}} z^2\nu_t(\mathrm{d}z)\,\mathrm{d}t$ with $\|z^2\star\nu\|_{L_\infty([0,T],\lambda)} < \infty$.

Eventually, plugging the estimates for $\mathrm{I}_{(5.27)}$–$\mathrm{IV}_{(5.27)}$ into (5.27), and using the fact that $D_{(5.26)} \leq D_{(5.24)}$, we derive a constant $c_{(5.28)} > 0$ independent of τ and ε such that

$$\begin{aligned}
&\|E^{\mathrm{D}}(\vartheta,\tau|\varepsilon,\kappa)\|_{\mathrm{BMO}_2^{\overline{\Phi}}(\mathbb{P})} \\
&\leq c_{(5.28)} D_{(5.26)} \\
&\leq c_{(5.28)} c_{(5.24)} \begin{cases} \|\tau\|_\theta & \text{if (3.17) holds,} \\ \left[1 + \log^+\left(\frac{1}{\varepsilon}\right) + \log^+\left(\frac{1}{\|\tau\|_\theta}\right)\right]\|\tau\|_\theta & \text{if (3.16) holds with } \alpha = 1, \\ \|\tau\|_\theta + \varepsilon^{1-\alpha}\|\tau\|_\theta^{1-\kappa(\alpha-1)} & \text{if (3.16) holds with } \alpha\in(1,2]. \end{cases}
\end{aligned} \tag{5.28}$$

Step 3. Combining (5.28) and (5.20) with (5.11) yields the conclusion, where we remark that the condition (3.16) with $\alpha \in (0,1)$ implies (3.17) due to Lemma 5.4 (with $\gamma = 1$). □

5.2.3. *Proof of Proposition* 3.15. We first consider the particular case when $\mathbb{Q}=\mathbb{P}$, $r=\infty$ and $q=2$. By Definition 3.9(1),

$$n+1=\#\tau_n\le \#\tau_n\sqcup\rho(\varepsilon_n,\kappa)\le n+1+\mathcal{N}_{(3.7)}(\varepsilon_n,\kappa),$$

and hence,

$$n+1\le \|\#\tau_n\sqcup\rho(\varepsilon_n,\kappa)\|_{L_2(\mathbb{P})}\le n+1+\|\mathcal{N}_{(3.7)}(\varepsilon_n,\kappa)\|_{L_2(\mathbb{P})}.$$

Since $\inf_{n\ge 1}\sqrt[\alpha]{n}\varepsilon_n>0$ by assumption, we derive from (5.8) that

$$c_{(5.8)}=T\|\mathbb{1}_{\{|z|>1\}}\star\nu\|_{L_\infty(\mathbb{P}\otimes\lambda)}+\varepsilon_n^{-\alpha}\sup_{r\in(0,1)}\|r^\alpha\mathbb{1}_{\{r<|z|\le 1\}}\star\nu\|_{L_\infty(\mathbb{P}\otimes\lambda)}\frac{T^{1-\alpha\kappa}}{1-\alpha\kappa}\le cn$$

for some constant $c>0$ independent of n. Using (5.8) gives the desired conclusion.

We next assume a probability measure $\mathbb{Q}\ll\mathbb{P}$ with $\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}\in L_r(\mathbb{P})$. Since $\frac{1}{2/q}+\frac{1}{r}=1$, applying Hölder's inequality yields

$$\|\#\tau_n\sqcup\rho(\varepsilon_n,\kappa)\|_{L_q(\mathbb{Q})}\le\|\#\tau_n\sqcup\rho(\varepsilon_n,\kappa)\|_{L_2(\mathbb{P})}\|\mathrm{d}\mathbb{Q}/\mathrm{d}\mathbb{P}\|^{1/q}_{L_r(\mathbb{P})},$$

and hence, (3.19) follows. □

5.2.4. *Verification of the assertion in Example* 3.19. We proceed in three steps.

Step 1. Set $\underline{X}_t:=\inf_{0\le s\le t}X_s$. We first prove that for any $0<\tilde{\varepsilon}<\frac12$ there exist a $\hat{\delta}^{\tilde{\varepsilon},K}:=\delta(\tilde{\varepsilon},K,\nu,\hat{\gamma})>0$ and a stopping time $\rho:=\rho(\tilde{\varepsilon},K)$ such that for any $0<\delta<\hat{\delta}^{\tilde{\varepsilon},K}$ the event

$$E^{\delta,\tilde{\varepsilon},K}_{\rho,\Delta X_\rho,X_{\rho-},\underline{X}_\rho}:=\left\{\rho\in(0,\delta],\,\Delta X_\rho\in(\tilde{\varepsilon},2\tilde{\varepsilon}),\,\frac{K}{2}\le X_{\rho-}\le\frac{5K}{2}+\frac{11}{2},\,\underline{X}_\rho\ge-\frac{K}{2}-1\right\}$$

has positive probability. Indeed, for $0<\tilde{\varepsilon}<\frac12$, we set $\hat{\gamma}^{(\tilde{\varepsilon})}:=\hat{\gamma}-\int_{\tilde{\varepsilon}<x<2\tilde{\varepsilon}}x\nu(\mathrm{d}x)$ and

$$X^{(\tilde{\varepsilon})}_t:=t\hat{\gamma}^{(\tilde{\varepsilon})}+\int_0^t\int_{\tilde{\varepsilon}<x<2\tilde{\varepsilon}}xN(\mathrm{d}s,\mathrm{d}x),\quad t\in[0,1].$$

Since $(0,1)\subset\operatorname{supp}(\nu)$ by assumption, it implies that $\nu((\tilde{\varepsilon},2\tilde{\varepsilon}))\in(0,\infty)$, and $X^{(\tilde{\varepsilon})}$ is hence a Poisson process with drift. Denote

$$\hat{X}^{(\tilde{\varepsilon})}_t:=X_t-X^{(\tilde{\varepsilon})}_t=\int_0^t\int_{\mathbb{R}_0\setminus(\tilde{\varepsilon},2\tilde{\varepsilon})}x\widetilde{N}(\mathrm{d}s,\mathrm{d}x),\quad t\in[0,1].$$

Remark that $X^{(\tilde{\varepsilon})}$ and $\hat{X}^{(\tilde{\varepsilon})}$ are independent Lévy processes. Let $\hat{\delta}^{\tilde{\varepsilon},K}_{(5.29)}>0$ be such that

$$\hat{\delta}^{\tilde{\varepsilon},K}_{(5.29)}|\hat{\gamma}^{(\tilde{\varepsilon})}|\le 1\quad\text{and}\quad\hat{\delta}^{\tilde{\varepsilon},K}_{(5.29)}\le\frac{K^2}{16}\left(\int_{\mathbb{R}}x^2\nu(\mathrm{d}x)\right)^{-1}. \tag{5.29}$$

We seek the desired stopping time ρ among jumping times of $X^{(\tilde{\varepsilon})}$. For any $\delta\in(0,\hat{\delta}^{\tilde{\varepsilon},K}_{(5.29)})$, since $X=X^{(\tilde{\varepsilon})}+\hat{X}^{(\tilde{\varepsilon})}$ and $\underline{X}\ge\underline{X}^{(\tilde{\varepsilon})}+\underline{\hat{X}}^{(\tilde{\varepsilon})}$, and notice that the event

$$\{\rho\in(0,\delta],\,\Delta X_\rho\in(\tilde{\varepsilon},2\tilde{\varepsilon})\}=\{\rho\in(0,\delta],\,\Delta X^{(\tilde{\varepsilon})}_\rho\in(\tilde{\varepsilon},2\tilde{\varepsilon})\}$$

is independent of $\hat{X}^{(\tilde{\varepsilon})}$, we get

$$\begin{aligned}&\mathbb{P}(E^{\delta,\tilde{\varepsilon},K}_{\rho,\Delta X_\rho,X_{\rho-},\underline{X}_\rho})\\&\quad\ge\mathbb{P}\left(\left\{\rho\in(0,\delta],\,\Delta X^{(\tilde{\varepsilon})}_\rho\in(\tilde{\varepsilon},2\tilde{\varepsilon}),\,K\le X^{(\tilde{\varepsilon})}_{\rho-}\le 2K+\frac{11}{2},\,\underline{X}^{(\tilde{\varepsilon})}_\rho\ge-1\right\}\right.\end{aligned}$$

$$\cap \left\{ -\frac{K}{2} \le \hat{X}^{(\tilde{\varepsilon})}_{\rho-} \le \frac{K}{2}, \underline{\hat{X}}^{(\tilde{\varepsilon})}_{\rho} \ge -\frac{K}{2} \right\} \Big)$$
$$\ge \mathbb{P}\Big(\rho \in (0,\delta], \Delta X^{(\tilde{\varepsilon})}_{\rho} \in (\tilde{\varepsilon}, 2\tilde{\varepsilon}), K \le X^{(\tilde{\varepsilon})}_{\rho-} \le 2K + \frac{11}{2}, \underline{X}^{(\tilde{\varepsilon})}_{\rho} \ge -1, \sup_{0\le t\le\delta} |\hat{X}^{(\tilde{\varepsilon})}_t| \le \frac{K}{2} \Big)$$
$$= \mathbb{P}\Big(\rho \in (0,\delta], \Delta X^{(\tilde{\varepsilon})}_{\rho} \in (\tilde{\varepsilon}, 2\tilde{\varepsilon}), K \le X^{(\tilde{\varepsilon})}_{\rho-} \le 2K + \frac{11}{2}, \underline{X}^{(\tilde{\varepsilon})}_{\rho} \ge -1 \Big)$$
$$\times \mathbb{P}\Big(\sup_{0\le t\le\delta} |\hat{X}^{(\tilde{\varepsilon})}_t| \le \frac{K}{2} \Big)$$
$$= \mathbb{P}\Big(\rho \in (0,\delta], \Delta X^{(\tilde{\varepsilon})}_{\rho} \in (\tilde{\varepsilon}, 2\tilde{\varepsilon}), K \le X^{(\tilde{\varepsilon})}_{\rho-} \le 2K + \frac{11}{2} \Big) \mathbb{P}\Big(\sup_{0\le t\le\delta} |\hat{X}^{(\tilde{\varepsilon})}_t| \le \frac{K}{2} \Big),$$

where we use the fact that $X^{(\tilde{\varepsilon})}_t \ge t\hat{\gamma}^{(\tilde{\varepsilon})} \ge -\hat{\delta}^{\tilde{\varepsilon},K}_{(5.29)} |\hat{\gamma}^{(\tilde{\varepsilon})}| \ge -1$ for all $t \in (0, \hat{\delta}^{\tilde{\varepsilon},K}_{(5.29)}]$ a.s. to get the last equality. For the second factor, applying Doob's maximal inequality yields

$$\mathbb{P}\Big(\sup_{0\le t\le\delta} |\hat{X}^{(\tilde{\varepsilon})}_t| \le \frac{K}{2} \Big) \ge 1 - \frac{16}{K^2} \mathbb{E}|\hat{X}^{(\tilde{\varepsilon})}_{\delta}|^2 = 1 - \frac{16\delta}{K^2} \int_{\mathbb{R}\setminus(\tilde{\varepsilon},2\tilde{\varepsilon})} x^2 \nu(\mathrm{d}x) > 0$$

for any $0 < \delta < \hat{\delta}^{\tilde{\varepsilon},K}_{(5.29)}$. Therefore, it remains to show the existence of a stopping time ρ with

$$\mathbb{P}\Big(\rho \in (0,\delta], \Delta X^{(\tilde{\varepsilon})}_{\rho} \in (\tilde{\varepsilon}, 2\tilde{\varepsilon}), K \le X^{(\tilde{\varepsilon})}_{\rho-} \le 2K + \frac{11}{2} \Big) > 0 \quad \forall \delta \in \big(0, \hat{\delta}^{\tilde{\varepsilon},K}_{(5.29)}\big).$$

Let $K_{\tilde{\varepsilon}} \in \mathbb{N}$ be such that

$$\tilde{\varepsilon}(K_{\tilde{\varepsilon}} - 2) \ge K + 1 \quad \text{and} \quad \tilde{\varepsilon}(K_{\tilde{\varepsilon}} - 3) < K + 1.$$

We define ρ to be the $K_{\tilde{\varepsilon}}$-th jump time of $X^{(\tilde{\varepsilon})}$,

$$\rho := \rho_{K_{\tilde{\varepsilon}}}. \tag{5.30}$$

Then it is clear that $\Delta X^{(\tilde{\varepsilon})}_{\rho_{K_{\tilde{\varepsilon}}}} \in (\tilde{\varepsilon}, 2\tilde{\varepsilon})$. On the set $\{\rho_{K_{\tilde{\varepsilon}}} \le \delta\}$ of positive probability one has

$$X^{(\tilde{\varepsilon})}_{\rho_{K_{\tilde{\varepsilon}}}-} \ge X^{(\tilde{\varepsilon})}_{\rho_{K_{\tilde{\varepsilon}}}} - 2\tilde{\varepsilon} \ge -\hat{\delta}^{\tilde{\varepsilon},K}_{(5.29)} |\hat{\gamma}^{(\tilde{\varepsilon})}| + \tilde{\varepsilon} K_{\tilde{\varepsilon}} - 2\tilde{\varepsilon} \ge K,$$
$$X^{(\tilde{\varepsilon})}_{\rho_{K_{\tilde{\varepsilon}}}-} \le X^{(\tilde{\varepsilon})}_{\rho_{K_{\tilde{\varepsilon}}}} - \tilde{\varepsilon} \le \hat{\delta}^{\tilde{\varepsilon},K}_{(5.29)} |\hat{\gamma}^{(\tilde{\varepsilon})}| + 2\tilde{\varepsilon} K_{\tilde{\varepsilon}} - \tilde{\varepsilon} \le 1 + 2K + 2 + 5\tilde{\varepsilon} \le 2K + \frac{11}{2},$$

which then verifies the assertion.

Step 2. Condition (3.23) is equivalent to $e^{\Delta X} \ge c_{(3.23)}$. Let $\hat{\delta}^{\tilde{\varepsilon},K}_{(5.31)}$ be such that

$$0 < \hat{\delta}^{\tilde{\varepsilon},K}_{(5.31)} \le T_0 \wedge \hat{\delta}^{\tilde{\varepsilon},K}_{(5.29)} \quad \text{and} \quad \sup_{0\le t\le \hat{\delta}^{\tilde{\varepsilon},K}_{(5.31)}} |\phi(t,1) - \phi(0,1)| \le \frac{1}{4} c_{(3.24)} K. \tag{5.31}$$

For $\rho = \rho_{K_{\tilde{\varepsilon}}}$ in (5.30) and for any $0 < \delta < \hat{\delta}^{\tilde{\varepsilon},K}_{(5.31)}$, on the set $E^{\delta,\tilde{\varepsilon},K}_{\rho,\Delta X_\rho, X_{\rho-}, \underline{X}_\rho}$ we have

$$\frac{S_{\rho-}}{\overline{\Phi}_\rho} |\phi(\rho, S_{\rho-}) - \phi(0,1)|$$
$$= \frac{c_{(3.23)} S_{\rho-}}{\Theta(\eta)_\rho S_\rho} |\phi(\rho, S_{\rho-}) - \phi(0,1)|$$
$$= \frac{c_{(3.23)}}{\Theta(\eta)_\rho e^{\Delta X_\rho}} |\phi(\rho, S_{\rho-}) - \phi(\rho,1) + \phi(\rho,1) - \phi(0,1)|$$

$$
\begin{aligned}
&\geq c_{(3.23)} \mathrm{e}^{(1-\eta)\underline{X}_\rho - 1} \big| |\phi(\rho, S_{\rho-}) - \phi(\rho, 1)| - |\phi(\rho, 1) - \phi(0, 1)| \big| \\
&\geq c_{(3.23)} \mathrm{e}^{(1-\eta)\underline{X}_\rho - 1} \Big[c_{(3.24)} |\mathrm{e}^{X_{\rho-}} - 1| - \sup_{0 \leq t \leq \delta} |\phi(t, 1) - \phi(0, 1)| \Big] \\
&\geq c_{(3.23)} \mathrm{e}^{(1-\eta)\underline{X}_\rho - 1} \Big[c_{(3.24)} X_{\rho-} - \sup_{0 \leq t \leq \delta} |\phi(t, 1) - \phi(0, 1)| \Big] \\
&\geq \frac{1}{4} c_{(3.23)} c_{(3.24)} K \mathrm{e}^{(1-\eta)\underline{X}_\rho - 1} \geq \frac{1}{4} c_{(3.23)} c_{(3.24)} \mathrm{e}^{-(1-\eta)(\frac{K}{2}+1)-1} K,
\end{aligned}
$$

where we use $\Theta(\eta)_\rho = \sup_{u \leq \rho} \mathrm{e}^{(\eta-1)X_u} \leq \mathrm{e}^{-(1-\eta)\underline{X}_\rho}$ and $\Delta X_\rho < 2\tilde{\varepsilon} < 1$ to obtain the first estimate.

Step 3. It is straightforward to check that conditions in Section 2.3 are satisfied, where we exploit the relation between the Lévy measures of X and of Z given in Appendix A to obtain the square integrability of Z, which verifies **[Z]**. We now use parameters in (3.25). For any $\varepsilon \in (0, 1)$, since $\Delta Z = \mathrm{e}^{\Delta X} - 1$ and $0 < \varepsilon r_* < 2^{-\kappa}/3 \leq 1/3$, one has

$$
\varepsilon r_* < |\Delta Z| < \varepsilon 3 r_* \;\Leftarrow\; \varepsilon r_* < \mathrm{e}^{\Delta X} - 1 < \varepsilon 3 r_* \;\Leftarrow\; \ln(1+\varepsilon r_*) < \Delta X < 2\ln(1+\varepsilon r_*).
$$

Then, for $\tilde{\varepsilon} := \ln(1+\varepsilon r_*) \in (0, \frac{1}{2})$, $\tau = (t_i)_{i=0}^n \in \mathcal{T}_{\det}$ with $\|\tau\|_1 \leq 1/2$, and $0 < \delta < t_1 \wedge \hat{\delta}^{\tilde{\varepsilon},K}_{(5.31)}$, it follows from *Step 2* that

$$
E_{(3.21)} \supseteq E^{\delta,\tilde{\varepsilon},K}_{\rho_*, \Delta X_{\rho_*}, X_{\rho_*-}, \underline{X}_{\rho_*}}.
$$

According to *Step 1*, we infer that $\mathbb{P}(E_{(3.21)}) > 0$. Eventually, the weight $\overline{\Phi} \in \mathcal{SM}_2(\mathbb{P})$ because of the assumption $\int_{|x|>1} \mathrm{e}^{2x} \nu(\mathrm{d}x) < \infty$ and Proposition B.2.

5.3. *Proofs of results in Section* 4.4.

5.3.1. *Proof of Proposition* 4.6. Note that Assumption 4.1 implies $\int_{|x|>1} \mathrm{e}^{2x} \nu(\mathrm{d}x) < \infty$. Since g has at most linear growth at infinity, one has $g(S_T) \in L_2(\mathbb{P})$.

(1) follows from Proposition B.2.

(2) We let $\ell := \nu$ in (C.1) and obtain from (4.2) that

$$
\vartheta^g_t = c^{-2}_{(4.2)} \Gamma_\nu(T-t, S_t) \quad \text{a.s.}, \forall t \in (0, T).
$$

Let us examine cases in Table 1. Using Proposition C.1(1) and (2) yields A1 and A2 respectively. For A3, since $\nu \in \mathscr{S}(\alpha)$, Remark 4.4(1) asserts that $\sup_{r \in (0,1)} r^\alpha \int_{r<|x|\leq 1} \nu(\mathrm{d}x) < \infty$. Applying Proposition C.1(3) with $\beta = \alpha$ yields

$$
|\vartheta^g_t| \leq c^{-2}_{(4.2)} c_{(\mathrm{C}.2)} U(t) S_t^{\eta-1} \text{ a.s.}, \forall t \in (0, T).
$$

The respective estimate for $|\vartheta^g_0|$ can be easily deduced by using the right continuity of ϑ^g, U and S.

(3) The SDE for $S = \mathrm{e}^X$ is

$$
\mathrm{d}S_t = S_{t-}\,\mathrm{d}Z_t, \quad S_0 = 1, \tag{5.32}
$$

where Z is *another* Lévy process under $\mathbb{P}$. Under Assumption 4.1, it is known that Z is also an $L_2(\mathbb{P})$-martingale with zero mean (see, e.g., [6], Proposition 8.20). Hence, conditions **[S]** and **[Z]** in Section 2.3 are fulfilled. Moreover, $\vartheta^g \in \mathcal{A}(S)$ due to Proposition 4.2(1, 2).

Let us now verify Assumption 3.4. Since $M = \vartheta^g S$ is an $L_2(\mathbb{P})$-martingale by Proposition 4.2(2), Assumption 3.1 holds because of Example 3.3 (with $V \equiv 0$). Thanks to (4.4), the growth condition (3.2) is satisfied for θ given in Table 1 case-wise. We now only need to check the curvature condition (3.3). If $U \equiv 1$ (in A1 with $\eta = 1$ and in A2), then the martingale M is closed in $L_2(\mathbb{P})$ by $M_T := L_2(\mathbb{P})\text{-}\lim_{t\uparrow T} M_t$ due to (4.4) and $\Phi(\eta) \in \mathcal{SM}_2(\mathbb{P})$. Then, for $\theta = 1$ and for any $a \in [0, T)$, one has, a.s.,

$$\begin{aligned}\mathbb{E}^{\mathcal{F}_a}\biggl[\int_{(a,T)} \Upsilon(\cdot, \mathrm{d}t)\biggr] &= \mathbb{E}^{\mathcal{F}_a}\biggl[\int_{(a,T)} \mathrm{d}\langle M\rangle_t + \int_{(a,T)} M_t^2\,\mathrm{d}t\biggr]\\ &\le \mathbb{E}^{\mathcal{F}_a}\Bigl[|M_T - M_a|^2 + c_{(4.4)}^2 (T-a) \sup_{t\in(a,T)} \Phi(\eta)_t^2\Bigr]\\ &\le c_{(4.4)}^2 (T+1) \|\Phi(\eta)\|_{\mathcal{SM}_2(\mathbb{P})}^2 \Phi(\eta)_a^2.\end{aligned}$$

For remaining cases, we set $\hat\theta := \eta$ in A1 for $\eta \in (0, 1)$, and set $\hat\theta := \frac{2(1+\eta)}{\alpha} - 1 \in (0, 1]$ in A3. Then, for any $\theta \in (0, \hat\theta)$, using the function U in Table 1 we get

$$(T-t)^{1-\theta} M_t^2 \le c_{(4.4)}^2 (T-t)^{1-\theta} U(t)^2 \Phi(\eta)_t^2 \to 0 \quad \text{a.s. as } t \uparrow T.$$

Thus, for any $a \in [0, T)$, a.s.,

$$\mathbb{E}^{\mathcal{F}_a}\biggl[\int_{(a,T)} (T-t)^{1-\theta} M_t^2\,\mathrm{d}t\biggr] \le c_{(5.33)} c_{(4.4)}^2 \|\Phi(\eta)\|_{\mathcal{SM}_2(\mathbb{P})}^2 \Phi(\eta)_a^2 \tag{5.33}$$

for some constant $c_{(5.33)} > 0$ depending at most on $\hat\theta, \theta, T$. Integrating by parts and applying conditional Itô's isometry yield, a.s.,

$$\begin{aligned}&\mathbb{E}^{\mathcal{F}_a}\biggl[\int_{(a,T)} (T-t)^{1-\theta}\,\mathrm{d}\langle M\rangle_t\biggr]\\ &\quad= \mathbb{E}^{\mathcal{F}_a}\biggl[\lim_{a<b\uparrow T} \int_{(a,b]} (T-t)^{1-\theta}\,\mathrm{d}\langle M\rangle_t\biggr]\\ &\quad= \mathbb{E}^{\mathcal{F}_a}\biggl[\lim_{a<b\uparrow T}\biggl[(T-b)^{1-\theta}|M_b - M_a|^2 + (1-\theta)\int_{(a,b]} (T-t)^{-\theta}|M_t - M_a|^2\,\mathrm{d}t\biggr]\biggr]\\ &\quad\le (1-\theta)\mathbb{E}^{\mathcal{F}_a}\biggl[\int_{(a,T)} (T-t)^{-\theta} M_t^2\,\mathrm{d}t\biggr] \le c_{(5.34)} c_{(4.4)}^2 \|\Phi(\eta)\|_{\mathcal{SM}_2(\mathbb{P})}^2 \Phi(\eta)_a^2\end{aligned} \tag{5.34}$$

for some $c_{(5.34)} = c_{(5.34)}(\hat\theta, \theta, T) > 0$. Combining (5.33) with (5.34) yields the desired conclusion.

For the particular case $\sigma = 0$, it is easy to check that Assumption 3.13 holds true. □

5.3.2. *Proof of Corollary* 4.7. Let ν_Z denote the Lévy measure of Z (which appears in the SDE (5.32)) under $\mathbb{P}$. By the relation between the Lévy measures of Z and of X given in Appendix A, some simple calculations yield that, for any $\alpha \in [0, 2]$,

$$\sup_{r\in(0,1)} r^\alpha \int_{r<|z|\le 1} \nu_Z(\mathrm{d}z) < \infty \quad \Leftrightarrow \quad \sup_{r\in(0,1)} r^\alpha \int_{r<|x|\le 1} \nu(\mathrm{d}x) < \infty.$$

(1) Since $\int_{|x|>1} \mathrm{e}^{2x} \nu(\mathrm{d}x) < \infty$ by Assumption 4.1, we get from Proposition B.2 that $\Phi(\eta)$ belongs to $\mathcal{SM}_2(\mathbb{P})$. Let us examine each case in Table 2.

<u>For B1</u>, the given range of (η, α) yields to $\int_{|x|\le 1} |x|^{1+\eta} \nu(\mathrm{d}x) < \infty$. We use Proposition 4.6(3), case A2 in Table 1, to obtain $\theta = 1$. Then, applying Theorem 3.16(3) gives the corresponding $R(n)$ and ε_n.

<u>For B2</u>, we note that $\mathscr{S}(\alpha) \subset \mathscr{U}\mathscr{S}(\alpha)$. Using Proposition 4.6(3), case A3 in Table 1, to get the range of θ, and then applying Theorem 3.16(3) we get the conclusions for $R(n)$ and ε_n.

For B3, we first get $\theta = 1$ due to Table 1, case A1. We decompose E^{adap} into three components E^{C}, E^{S} and E^{D} as in (5.10). For the "continuous part" E^{C}, (5.12) yields

$$\sup_{n\geq 1} \sqrt{n} \left\| E^{\mathrm{C}}\left(\vartheta^g, \tau_n^1 \,\middle|\, \sqrt{1/n}, 0\right)\right\|_{\mathrm{BMO}_2^{\overline{\Phi}(1)}(\mathbb{P})} < \infty.$$

For the "small jump part" E^{S} and the "drift part" E^{D}, we apply (5.20) and (5.28) with $\alpha = 2$ respectively to obtain

$$\sup_{n\geq 1} \sqrt{n}\left(\left\| E^{\mathrm{S}}\left(\vartheta^g, \tau_n^1 \,\middle|\, \sqrt{1/n}, 0\right)\right\|_{\mathrm{BMO}_2^{\overline{\Phi}(1)}(\mathbb{P})} + \left\| E^{\mathrm{D}}\left(\vartheta^g, \tau_n^1 \,\middle|\, \sqrt{1/n}, 0\right)\right\|_{\mathrm{BMO}_2^{\overline{\Phi}(1)}(\mathbb{P})}\right) < \infty.$$

Thus, the desired assertion follows from (5.11).

For B4, it is similar to B3.

(2) Since $\int_{|x|>1} \mathrm{e}^{px} \nu(\mathrm{d}x) < \infty$, it follows from Propositions B.2 and B.1(2) that $\Phi(\eta)$ and $\overline{\Phi}(\eta)$ belong to $\mathcal{SM}_p(\mathbb{P})$ for all $\eta \in [0, 1]$. Hence, the conclusion follows from Proposition 2.5(1, 2). □

APPENDIX A: EXPONENTIAL LÉVY PROCESSES

Let X be a Lévy process with characteristic triplet (γ, σ, ν) as in Section 4.1. Then, the ordinary exponential $S = \mathrm{e}^X$ can be represented as the *Doléans–Dade exponential* $\mathcal{E}(Z)$ of another Lévy process Z (see, e.g., [1], Theorem 5.1.6), that is, $\mathrm{e}^X = \mathcal{E}(Z)$ for the process Z in the SDE (5.32). Z is also known as the stochastic logarithm of S, that is, $Z = \mathcal{L}(S)$.

The path relation of X and Z is given by

$$Z_t = X_t + \frac{\sigma^2 t}{2} + \sum_{s\in[0,t]} \left(\mathrm{e}^{\Delta X_s} - 1 - \Delta X_s\right), \quad t \in [0, T] \text{ a.s.},$$

which then implies $\Delta Z = \mathrm{e}^{\Delta X} - 1$. The relation between the characteristic (γ, σ, ν) of X and $(\gamma_Z, \sigma_Z, \nu_Z)$ of Z is provided, for example, in [1], Theorem 5.1.6. In particular, one has $\sigma_Z = \sigma$ and $\nu_Z(\cdot) = \int_{\mathbb{R}} \mathbb{1}_{\{\mathrm{e}^x - 1 \in \cdot\}} \nu(\mathrm{d}x)$.

APPENDIX B: REGULARITY OF WEIGHT PROCESSES

We recall $\overline{\Phi}$ from (3.9) and $\mathcal{SM}_p(\mathbb{P})$ from Definition 2.1.

PROPOSITION B.1.

(1) *Let $p, q, r \in (0, \infty)$ with $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. Then, for any $\Phi, \Psi \in \mathrm{CL}^+([0, T])$ one has $\|\Phi\Psi\|_{\mathcal{SM}_r(\mathbb{P})} \leq \|\Phi\|_{\mathcal{SM}_p(\mathbb{P})}\|\Psi\|_{\mathcal{SM}_q(\mathbb{P})}$.*

(2) *If $\Phi \in \mathcal{SM}_p(\mathbb{P})$ for some $p \in (0, \infty)$, then $\overline{\Phi} \in \mathcal{SM}_p(\mathbb{P})$.*

PROOF. Item (1) is provided in [17], Proposition A.2, and Item (2) in the supplementary material [37], subsection F.1. □

Let $X = (X_t)_{t\in[0,T]}$ be a Lévy process with characteristic triplet (γ, σ, ν) and exponent ψ as in Section 4.1. Recall $S = \mathrm{e}^X$ and $\Phi(\eta)$ from (4.3).

PROPOSITION B.2. *If $\int_{|x|>1} \mathrm{e}^{qx} \nu(\mathrm{d}x) < \infty$ for some $q \in (1, \infty)$, then $\Phi(\eta) \in \mathcal{SM}_q(\mathbb{P})$ for all $\eta \in [0, 1]$. Moreover, for $c_q := (\frac{q}{q-1})^q$ one has*

$$\|\Phi(\eta)\|_{\mathcal{SM}_q(\mathbb{P})}^q \leq \mathrm{e}^{T|\psi(-\mathrm{i})|(2q+1)} 2^{1-\eta} c_q^2 \|S_T\|_{L_q(\mathbb{P})}^q.$$

PROOF. The proof is given in the Supplementary Material [37], subsection F.2. □

APPENDIX C: GRADIENT TYPE ESTIMATES FOR A LÉVY SEMIGROUP ON HÖLDER SPACES

Let X be a Lévy process with characteristic triplet (γ, σ, ν) as in Section 4.1. Let $\eta \in [0,1]$ and recall $C^{0,\eta}(\mathbb{R}_+)$ from Definition 4.3(1). Assume $\int_{|x|>1} \mathrm{e}^{\eta x}\nu(\mathrm{d}x) < \infty$. We define the map $P_t \colon C^{0,\eta}(\mathbb{R}_+) \to C^{0,\eta}(\mathbb{R}_+)$ by $P_t g(y) := \mathbb{E} g(y\mathrm{e}^{X_t})$, $y > 0, t \geq 0$.

Motivated by formula (4.2), for a Lévy measure ℓ and $g \in C^{0,\eta}(\mathbb{R}_+)$, we formally set

$$\Gamma_\ell(t, y) := \sigma^2 \partial_y P_t g(y) + \int_{\mathbb{R}} \frac{P_t g(\mathrm{e}^x y) - P_t g(y)}{y} (\mathrm{e}^x - 1)\ell(\mathrm{d}x), \quad t, y > 0, \tag{C.1}$$

where $\partial_y P_t g(y) := 0$ if $\sigma = 0$. Although we choose $\ell = \nu$ for (4.2), it is useful to consider the general ℓ because it might have applications in other contexts (e.g., see [36]).

Proposition C.1(3) below is a variant of [17], Theorem 9.18, in the exponential Lévy setting. Here, the exponent of the time variable t in the obtained estimates is the same as in [17], Theorem 9.18. We recall $\mathscr{S}(\alpha)$ from Definition 4.3(2).

PROPOSITION C.1. *Let ℓ be a Lévy measure and $g \in C^{0,\eta}(\mathbb{R}_+)$ with $\eta \in [0,1]$. Assume that $\int_{|x|>1} \mathrm{e}^{(\eta+1)x}\ell(\mathrm{d}x) < \infty$. Then, for any $T \in (0,\infty)$ there is a $c_{(C.2)} > 0$ such that*

$$|\Gamma_\ell(t, y)| \leq c_{(C.2)} V(t) y^{\eta - 1} \quad \forall (t, y) \in (0, T] \times \mathbb{R}_+, \tag{C.2}$$

where the cases for $V(t)$ are provided as follows:

(1) *If $\sigma > 0$ and $\int_{|x|>1} \mathrm{e}^{2x}\nu(\mathrm{d}x) < \infty$, then $V(t) = t^{\frac{\eta-1}{2}}$.*
(2) *If $\sigma = 0$, $\int_{|x|>1} \mathrm{e}^{\eta x}\nu(\mathrm{d}x) < \infty$ and $\int_{|x|\leq 1} |x|^{\eta+1}\ell(\mathrm{d}x) < \infty$, then $V(t) = 1$.*
(3) *If $\sigma = 0$, $\eta \in [0,1)$ and if the following two conditions hold*:
    (a) *$\nu \in \mathscr{S}(\alpha)$ for some $\alpha \in (0,2)$ and $\int_{|x|>1} \mathrm{e}^{x}\nu(\mathrm{d}x) < \infty$,*
    (b) *there is a $\beta \in [0,2]$ such that*

$$0 < c_{(C.3)} := \sup_{r \in (0,1)} r^\beta \int_{r<|x|\leq 1} \ell(\mathrm{d}x) < \infty, \tag{C.3}$$

then $V(t) = t^{\frac{\eta+1-\beta}{\alpha}}$ if $\beta \in (1+\eta, 2]$, $V(t) = \max\{1, \log(1/t)\}$ if $\beta = 1+\eta$, and $V(t) = 1$ if $\beta \in [0, 1+\eta)$.

Here, the constant $c_{(C.2)}$ may depend on β in Item (3).

PROOF. See the Supplementary Material [37], subsection G.2. □

Acknowledgments. A major part of this work was done when the author was affiliated to the Department of Mathematics and Statistics, University of Jyväskylä, Finland. The author would like to express his gratitude to Hannah Geiss and Stefan Geiss for fruitful discussions. He is also grateful to the Editors and the anonymous Reviewers for their helpful suggestions and constructive recommendations which led to a significant improvement of the article.

The author is also affiliated to the Department of Mathematics, Vinh University, 182 Le Duan, Vinh, Nghe An, Viet Nam (e-mail: thuannt@vinhuni.edu.vn).

This article is dedicated to Professor Phan Đức Thành on the occasion of his 86th birthday.

Funding. The author was mainly supported by the Project 298641 "*Stochastic Analysis and Nonlinear Partial Differential Equations, Interactions and Applications*" of the Academy of Finland.

SUPPLEMENTARY MATERIAL

Supplementary material for "Approximation of stochastic integrals with jumps via weighted BMO approach" by N.T. Thuan (DOI: 10.1214/24-AAP2075SUPP; .pdf). This document contains the proofs of Example 3.3, Lemmas 5.1, 5.2 and 5.4, Propositions 4.2, B.1(2), B.2 and C.1.

REFERENCES

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

[2] BENTH, F. E., DI NUNNO, G., LØKKA, A., ØKSENDAL, B. and PROSKE, F. (2003). Explicit representation of the minimal variance portfolio in markets driven by Lévy processes. *Math. Finance* **13** 55–72. Conference on Applications of Malliavin Calculus in Finance (Rocquencourt, 2001). MR1968096 https://doi.org/10.1111/1467-9965.t01-1-00005

[3] BLUMENTHAL, R. M. and GETOOR, R. K. (1961). Sample functions of stochastic processes with stationary independent increments. *J. Math. Mech.* **10** 493–516. MR0123362

[4] BRODÉN, M. and TANKOV, P. (2011). Tracking errors from discrete hedging in exponential Lévy models. *Int. J. Theor. Appl. Finance* **14** 803–837. MR2845218 https://doi.org/10.1142/S0219024911006760

[5] CHOULLI, T., KRAWCZYK, L. and STRICKER, C. (1998). $\mathscr{E}$-martingales and their applications in mathematical finance. *Ann. Probab.* **26** 853–876. MR1626523 https://doi.org/10.1214/aop/1022855653

[6] CONT, R. and TANKOV, P. (2004). *Financial Modelling with Jump Processes*. *Chapman & Hall/CRC Financial Mathematics Series*. CRC Press, Boca Raton, FL. MR2042661

[7] CONT, R., TANKOV, P. and VOLTCHKOVA, E. (2007). Hedging with options in models with jumps. In *Stochastic Analysis and Applications*. *Abel Symp.* **2** 197–217. Springer, Berlin. MR2397788 https://doi.org/10.1007/978-3-540-70847-6_8

[8] DELBAEN, F., MONAT, P., SCHACHERMAYER, W., SCHWEIZER, M. and STRICKER, C. (1997). Weighted norm inequalities and hedging in incomplete markets. *Finance Stoch.* **1** 181–227.

[9] DELLACHERIE, C. and MEYER, P.-A. (1982). *Probabilities and Potential. B: Theory of Martingales*. *North-Holland Mathematics Studies* **72**. North-Holland, Amsterdam. Translated from the French by J. P. Wilson. MR0745449

[10] DEREICH, S. and HEIDENREICH, F. (2011). A multilevel Monte Carlo algorithm for Lévy-driven stochastic differential equations. *Stochastic Process. Appl.* **121** 1565–1587. MR2802466 https://doi.org/10.1016/j.spa.2011.03.015

[11] FUKASAWA, M. (2011). Asymptotically efficient discrete hedging. In *Stochastic Analysis with Financial Applications*. *Progress in Probability* **65** 331–346. Birkhäuser, Basel. MR3050797 https://doi.org/10.1007/978-3-0348-0097-6_19

[12] FUKASAWA, M. (2014). Efficient discretization of stochastic integrals. *Finance Stoch.* **18** 175–208. MR3146491 https://doi.org/10.1007/s00780-013-0215-6

[13] GEISS, C. and GEISS, S. (2004). On approximation of a class of stochastic integrals and interpolation. *Stoch. Stoch. Rep.* **76** 339–362. MR2075477 https://doi.org/10.1080/10451120410001728445

[14] GEISS, C., GEISS, S. and LAUKKARINEN, E. (2013). A note on Malliavin fractional smoothness for Lévy processes and approximation. *Potential Anal.* **39** 203–230. MR3102985 https://doi.org/10.1007/s11118-012-9326-5

[15] GEISS, S. (2002). Quantitative approximation of certain stochastic integrals. *Stoch. Stoch. Rep.* **73** 241–270. MR1932161 https://doi.org/10.1080/1045112021000025934

[16] GEISS, S. (2005). Weighted BMO and discrete time hedging within the Black–Scholes model. *Probab. Theory Related Fields* **132** 13–38. MR2136865 https://doi.org/10.1007/s00440-004-0389-0

[17] GEISS, S. and THUAN, N. T. (2021). On Riemann–Liouville operators, BMO, gradient estimates in the Lévy–Itô space, and approximation. Available at arXiv:2009.00899v2.

[18] GEISS, S. and TOIVOLA, A. (2009). Weak convergence of error processes in discretizations of stochastic integrals and Besov spaces. *Bernoulli* **15** 925–954. MR2597578 https://doi.org/10.3150/09-BEJ197

[19] GEISS, S. and TOIVOLA, A. (2015). On fractional smoothness and L_p-approximation on the Gaussian space. *Ann. Probab.* **43** 605–638. MR3306001 https://doi.org/10.1214/13-AOP884

[20] GEISS, S. and YLINEN, J. (2020). Weighted bounded mean oscillation applied to backward stochastic differential equations. *Stochastic Process. Appl.* **130** 3711–3752. MR4092418 https://doi.org/10.1016/j.spa.2019.10.007

[21] GOBET, E. and LANDON, N. (2014). Almost sure optimal hedging strategy. *Ann. Appl. Probab.* **24** 1652–1690. MR3211007 https://doi.org/10.1214/13-AAP959

[22] GOBET, E. and TEMAM, E. (2001). Discrete time hedging errors for options with irregular payoffs. *Finance Stoch.* **5** 357–367. MR1849426 https://doi.org/10.1007/PL00013539

[23] HUBALEK, F., KALLSEN, J. and KRAWCZYK, L. (2006). Variance-optimal hedging for processes with stationary independent increments. *Ann. Appl. Probab.* **16** 853–885. MR2244435 https://doi.org/10.1214/105051606000000178

[24] JACOD, J., MÉLÉARD, S. and PROTTER, P. (2000). Explicit form and robustness of martingale representations. *Ann. Probab.* **28** 1747–1780. MR1813842 https://doi.org/10.1214/aop/1019160506

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

[26] KOHATSU-HIGA, A. and TANKOV, P. (2010). Jump-adapted discretization schemes for Lévy-driven SDEs. *Stochastic Process. Appl.* **120** 2258–2285. MR2684745 https://doi.org/10.1016/j.spa.2010.07.001

[27] MIKULEVICIUS, R. (2012). On the rate of convergence of simple and jump-adapted weak Euler schemes for Lévy driven SDEs. *Stochastic Process. Appl.* **122** 2730–2757. MR2926173 https://doi.org/10.1016/j.spa.2012.04.013

[28] PLATEN, E. and BRUTI-LIBERATI, N. (2010). *Numerical Solution of Stochastic Differential Equations with Jumps in Finance*. *Stochastic Modelling and Applied Probability* **64**. Springer, Berlin. MR2723480 https://doi.org/10.1007/978-3-642-13694-8

[29] PROTTER, P. E. (2004). *Stochastic Integration and Differential Equations*: *Stochastic Modelling and Applied Probability*, 2nd ed. *Applications of Mathematics* (*New York*) **21**. Springer, Berlin. MR2020294

[30] ROSENBAUM, M. and TANKOV, P. (2014). Asymptotically optimal discretization of hedging strategies with jumps. *Ann. Appl. Probab.* **24** 1002–1048. MR3199979 https://doi.org/10.1214/13-AAP940

[31] SATO, K. (2013). *Lévy Processes and Infinitely Divisible Distributions*. *Cambridge Studies in Advanced Mathematics* **68**. Cambridge Univ. Press, Cambridge. Translated from the 1990 Japanese original, Revised edition of the 1999 English translation. MR3185174

[32] SCHOUTENS, W. (2003). *Lévy Processes in Finance*: *Pricing Financial Derivatives*. Wiley, New York.

[33] SCHWEIZER, M. (2001). A guided tour through quadratic hedging approaches. In *Option Pricing*, *Interest Rates and Risk Management*. *Handb. Math. Finance* 538–574. Cambridge Univ. Press, Cambridge. MR1848562 https://doi.org/10.1017/CBO9780511569708.016

[34] TANKOV, P. (2011). Pricing and hedging in exponential Lévy models: Review of recent results. In *Paris–Princeton Lectures on Mathematical Finance* 2010. *Lecture Notes in Math.* **2003** 319–359. Springer, Berlin. MR2762364 https://doi.org/10.1007/978-3-642-14660-2_5

[35] TANKOV, P. and VOLTCHKOVA, E. (2009). Asymptotic analysis of hedging errors in models with jumps. *Stochastic Process. Appl.* **119** 2004–2027. MR2519354 https://doi.org/10.1016/j.spa.2008.10.002

[36] THUAN, N. T. (2022). Explicit Föllmer–Schweizer decomposition and discretization with jump correction in exponential Lévy models. Available at arXiv:2009.04328v2.

[37] THUAN, N. T. (2024). Supplement to "Approximation of stochastic integrals with jumps via weighted BMO approach." https://doi.org/10.1214/24-AAP2075SUPP